



Matematické základy digitálního zpracování obrazu

Eduard Sojka, Jan Gaura,
Michal Krumník

Text byl vytvořen v rámci realizace projektu *Matematika pro inženýry 21. století* (reg. č. CZ.1.07/2.2.00/07.0332), na kterém se společně podílela Vysoká škola báňská – Technická univerzita Ostrava a Západočeská univerzita v Plzni



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Eduard Sojka, Jan Gaura, Michal Krumnikl
Matematické základy digitálního zpracování obrazu

© Eduard Sojka, Jan Gaura, Michal Krumnikl, 2011
ISBN

Předmluva

Digitální zpracování a analýza obrazů jsou moderní a rychle se rozvíjející disciplíny. Jejich výsledky jsou dnes již běžně používány při úpravách a kompresi obrazů a videosekvencí, při rozpoznávání textů, v kontrolních a střežících systémech, v průmyslových robotech, ve fotogrammetrii i jinde. Dostatečný výkon a technologická úroveň současných výpočetních systémů umožnila pronikání digitálního zpracování obrazu dokonce i do spotřební elektroniky. Příkladem je digitální fotografie, digitální video a digitální televize.

Na Fakultě elektrotechniky a informatiky VŠB-TU Ostrava je ve studijním oboru „informatika a výpočetní technika“ problematika digitálního zpracování a analýzy obrazu přednášena v předmětech „digitální zpracování obrazu“ a „analýza obrazu“. Předkládaný učební text je určen pro potřeby obou těchto předmětů. Vychází z mnohaleté zkušenosti s výukou těchto předmětů a pokrývá značnou část látky, která je přednášena zejména v předmětu „digitální zpracování obrazu“. Poněkud teoretická povaha textu je kompenzována ve cvičeních, která jsou naopak věnována implementačním a jiným praktickým pohledům na probíranou tematiku, jakož i ukázkám v praxi používaných programových systémů. Jsme si vědomi toho, že i při velké péči, kterou jsme vypracování textu věnovali, se v něm mohou vyskytnout nejasnosti, překlepy nebo i chyby. Čtenářům budeme proto vděčni za upozornění na veškerá případná nedopatření i za připomínky a komentáře.

V Ostravě 6. června 2012

Autoři

Obsah

Předmluva	iii
1 Prostor obrazových signálů	1
1.1 Prostor signálů	1
1.1.1 Linearita prostoru signálů, báze v prostoru signálů	2
1.1.2 Skalární součin nad prostorem signálů	4
1.1.3 Re prezentace signálu pomocí bázevých funkcí	6
1.2 Operace nad prostorem signálů	8
1.2.1 Lineární operace	8
1.2.2 Operace invariantní vůči posuvu	8
1.2.3 Diracův impulz	9
1.2.4 Lineární sumace bodových zdrojů, konvoluce	11
1.2.5 Diskrétní konvoluce	14
2 Transformace obrazových signálů	16
2.1 Báze δ	16
2.2 Transformace signálu	17
2.3 Fourierova transformace	18
2.3.1 Spojitá Fourierova transformace	19
2.3.2 Vlastnosti spojité Fourierovy transformace	24
2.3.3 Diskrétní Fourierova transformace	27
2.3.4 Vlastnosti diskrétní Fourierovy transformace	28
2.4 Kosinová transformace	31
3 Stochastický přístup k popisu obrazových signálů	33
3.1 Náhodné pole	33
3.2 Momenty náhodného pole	34
3.3 Homogenní a ergodické náhodné pole	35
3.4 Výkonová spektrální hustota	36
3.5 Lineární operace nad náhodným polem	37
3.6 Statistická analýza obrazových transformací	39

4	Digitalizace obrazových signálů	41
4.1	Vzorkování deterministicky popsaných obrazů	41
4.2	Zpřesněné modely vzorkování	45
4.3	Aliasing	47
4.4	Praktické metody rekonstrukce obrazu	49
4.5	Kvantování obrazových signálů	54
5	Filtrace obrazových signálů	57
5.1	Nerekurzivní filtrace	57
5.1.1	Realizace nerekurzivního filtru v prostorové a frekvenční doméně	58
5.1.2	Filtry s nulovou fází	58
5.1.3	Návrh filtru s využitím výřezové funkce	59
5.1.4	Návrh optimálního filtru	60
5.2	Rekurzivní filtrace	62
5.2.1	Realizace rekurzivního filtru přímým výpočtem	62
5.2.2	Realizace rekurzivního filtru iteračním výpočtem	64
5.2.3	Návrh rekurzivního filtru	65
5.3	Obecný model degradace a rekonstrukce obrazu	66
5.4	Inverzní filtrace	67
5.5	Wienerův filtr ve frekvenční doméně	67
6	Bodové, algebraické a geometrické operace s obrazy	69
6.1	Bodové operace	69
6.1.1	Vliv bodových operací na histogram jasu	70
6.1.2	Vyrovnání histogramu	73
6.2	Algebraické operace s obrazy	74
6.2.1	Vliv algebraických operací na histogram jasu	75
6.2.2	Využití průměrování k redukci šumu	75
6.3	Geometrické transformace obrazu	76
7	Segmentace obrazu	84
7.1	Detekce hran	84
7.1.1	Gradientní metody hledání hran	85
7.1.2	Detekce hran hledáním průchodu druhé derivace nulou	88
7.2	Spojování hran	92
7.2.1	Houghova transformace	92
7.2.2	Cannyho detektor hran	94
7.2.3	Parametrické modely hrany	100
7.3	Detekce oblastí	102
7.3.1	Prahování	102
7.4	Segmentace jako úloha o minimalizaci funkcionálu	105

7.4.1	Aktivní kontury „bez hran“	106
7.5	Zpracování binárních obrazů	108
7.5.1	Matematická morfologie	108
8	Zpětná stereoprojekce	111
8.1	Model kamery	111
8.2	Dvojice kamer s rovnoběžnými optickými osami	114
8.3	Absolutní kalibrace a rekonstrukce	115
8.4	Relativní kalibrace a rekonstrukce	117
8.5	Epipolára	122
8.6	Automatizované hledání korespondence	123
8.6.1	Okénkové algoritmy	125
8.6.2	Dynamické programování	126
8.6.3	Globální metody	129
	Literatura	131
	Rejstřík	133

Kapitola 1

Prostor obrazových signálů

V této kapitole bude podán přehled základních teoretických nástrojů, které jsou potřebné pro kvalifikované zvládnutí problematiky digitálního zpracování obrazu. Čtenáři zde prosíme o určitou dávku trpělivosti. Bez pochopení skutečností uvedených v této kapitole se zpracováním obrazu nelze seriózně zabývat. Protože každý optický obraz je signálem, budeme v této kapitole poněkud obecněji a místo o obrazu nebo obrazovém signálu budeme často obecně hovořit pouze o signálu. Závěry, které zde uvedeme, platí pro jakékoli signály, a tedy i pro signály obrazové. Obrazové signály jsou zpravidla dvojrozměrné. S ohledem na tuto skutečnost budeme vztahy, které budou později prakticky využívány, ihned formulovat také v jejich dvojrozměrných variantách.

1.1 Prostor signálů

Předpokládáme, že *signál* je definován nad jistou podmnožinou (označme ji Ω) m -rozměrného euklidovského prostoru E^m (euklidovský prostor je prostor splňující Euklidovy postuláty; jedná se o běžný metrický prostor, s nímž jsme zvyklí pracovat např. v geometrii). Touto podmnožinou může být nějaká ohraničená nebo neohraničená souvislá oblast v E^m , množina jednotlivých bodů z E^m , ale také celý prostor E^m . Jestliže je oblast, nad níž je signál definován, souvislá, jedná se o tzv. *spojitý signál*. V dalším textu budeme *souřadnice bodů* takové oblasti označovat (x, y) , (u, v) atd. Jestliže je obraz definován pouze nad jednotlivými body, jedná se o tzv. *diskrétní signál*. Velmi často se zde používá uspořádání bodů v pravidelné (ekvidistantní) mřížce. Bez újmy na obecnosti lze pak předpokládat, že souřadnice bodů jsou celočíselné. V dalším textu budeme souřadnice bodů, nad nimiž je diskrétní obraz definován, značit jako (m, n) , (k, l) atd. Při praktické implementaci na počítači se používá diskrétní reprezentace. Spojitá reprezentace je však také užitečná. Využijeme ji zejména při teoretickém odvozování. Ve spojitém i diskrétním případě mohou mít obrazy buď konečnou a nebo i nekonečnou velikost. Představě o obrazu odpovídá samozřejmě více velikost konečná (nejčastěji si představujeme, že je obraz definován nad oblastí ve tvaru obdélníka). Nekonečnou velikost je možné považovat

za zjednodušení užitečné při odvozování, které může být někdy snazší pro velikost nekonečnou než konečnou.

Matematickou *reprezentací signálu* je funkce $f : \Omega \rightarrow \gamma$, kde γ je obor hodnot signálu. Množina γ může být např. množinou reálných nebo celých čísel (obrazy ve stupních šedi), množinou komplexních čísel (obrazy vznikající některými transformacemi), množinou vektorů o třech složkách (v barevných obrazech bývá často barva každého bodu popsána trojicí R, G, B , jejíž jednotlivé členy udávají intenzitu červené, zelené a modré barevné složky) nebo podmnožinou některé z uvedených množin. O signálu f řekneme, že je m -rozměrným signálem. Množinu funkcí realizujících zobrazení $\Omega \rightarrow \gamma$ nazveme m -rozměrným *signálovým prostorem* \mathcal{S} . Je tedy $\mathcal{S} = \{f | f : \Omega \rightarrow \gamma\}$. Úvahy provedené dále v této kapitole se budou týkat zejména signálů reálných a komplexních. Doplnující informace, které se týkají signálů reprezentovaných celými čísly a vektory, uvedeme na vhodném místě později.

Příklad 1.1. Mějme obraz v úrovních jasu rozměru 3×3 pixely. Hodnoty jasu v každém pixelu mohou být z rozsahu $0, \dots, 9$. Určete, kolik prvků má v daném případě prostor obrazových signálů \mathcal{S} .

Řešení. S uvážením podobnosti s dekadickými čísly o devíti řádech snadno docházíme k závěru, že hledaný počet prvků je 10^9 . ▲

1.1.1 Linearita prostoru signálů, báze v prostoru signálů

Při zpracování signálů dosti často předpokládáme, že pro signály platí *princip superpozice*. Matematickým vyjádřením této představy je, že signálový prostor předpokládáme lineární. Zopakujme, že v lineárním prostoru je definována operace sčítání prvků a násobení prvku konstantou (konstantou může být reálné nebo komplexní číslo). Připomeňme dále, že v lineárním prostoru existuje nulový prvek a ke každému prvku existuje prvek opačný.

Ve zpracování signálu a obrazu hraje významnou úlohu *báze* v prostoru signálů, což je pojem, který v dalším textu také připomeneme. Uvažujme nejprve prostor jednorozměrných diskrétních signálů (označme jej \mathcal{S}), kde definičním oborem Ω je množina M bodů. Každý bod je popsán svým indexem m . Bez újmy na obecnosti lze předpokládat, že index nabývá hodnot $0, 1, \dots, M - 1$. V prostoru \mathcal{S} uvažujme množinu $\{\varphi_k(m) | k = 0, 1, \dots, N - 1\}$ obsahující N diskrétních funkcí definovaných nad Ω . Funkce této množiny nazveme lineárně závislými, jestliže lze najít koeficienty a_k (z nichž alespoň jeden je různý od nuly) tak, že platí vztah

$$\sum_{k=0}^{N-1} a_k \varphi_k(m) = 0. \quad (1.1)$$

V opačném případě funkce nazveme lineárně nezávislými.

Množina $\varphi_k(m)$ je úplná v daném prostoru signálů, jestliže je možné každou funkci z tohoto prostoru vyjádřit jako lineární kombinaci funkcí z této množiny.

Lineárně nezávislý systém funkcí $\varphi_k(m)$, který je úplný, nazveme bází prostoru \mathcal{S} . Každou funkci $f(m)$ prostoru \mathcal{S} lze vyjádřit jako lineární kombinaci bázových funkcí

$$f(m) = \sum_{k=0}^{N-1} F_k \varphi_k(m) = \sum_{k=0}^{N-1} F(k) \varphi_k(m), \quad (1.2)$$

kde F_k jsou koeficienty, které obecně mohou být komplexní. Protože na \mathcal{S} můžeme pohlížet jako na prostor M -rozměrných vektorů, je zřejmé, že bázových funkcí je právě M . Pro zvolený systém bázových funkcí lze proto každý signál $f(m)$ prostoru \mathcal{S} reprezentovat uspořádanou M -ticí $(F_0, F_1, \dots, F_{M-1})$ koeficientů. Na tuto M -tici můžeme ovšem také pohlížet jako na diskretní funkci $F(k)$, $k = 0, 1, \dots, M-1$, čehož jsme již ve vztahu (1.2) využili.

Analogicky lze postupovat také pro jednorozměrné spojité signály. V tomto případě je definičním oborem signálů nespočetná množina $\Omega \subseteq \mathbb{E}^1$, a proto lze očekávat, že i báze bude obecně tvořena nespočetnou množinou funkcí. Takovou množinu můžeme formálně zapsat jako jedinou funkci $\varphi(u, x)$ dvou argumentů, kde $x \in \Omega$ a $u \in U$, $U \subseteq \mathbb{R}$ (\mathbb{R} značí množinu reálných čísel). Libovolný spojitý signál $f(x)$ lze pak vyjádřit jako lineární kombinaci pomocí vztahu

$$f(x) = \int_U F(u) \varphi(u, x) du. \quad (1.3)$$

S ohledem na zvolenou bázi $\varphi(u, x)$ lze tedy každý prvek $f(x)$ prostoru spojitých signálů reprezentovat pomocí funkce $F(u)$.

Závěry zatím uvedené snadno zobecníme také na signály dvojrozměrné. Předpokládáme, že v případě dvojrozměrného diskretního signálu je $\Omega = \{(m, n) | m = 0, 1, \dots, M-1; n = 0, 1, \dots, N-1\}$. Celkový počet bázových funkcí je proto $M \times N$. V případě dvojrozměrného spojitého signálu je podle očekávání poloha každého bodu popsána dvojicí souřadnic x, y a dále je $U \subseteq \mathbb{R} \times \mathbb{R}$. Pro dvojrozměrný diskretní a dvojrozměrný spojitý signál tak postupně máme

$$f(m, n) = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} F(k, l) \varphi_{k,l}(m, n), \quad (1.4)$$

$$f(x, y) = \iint_U F(u, v) \varphi(u, v, x, y) du dv. \quad (1.5)$$

Jak je zřejmé ze vztahu (1.4), nahlížíme ve shodě s poznámkou ke vztahu (1.2) v diskretním případě na koeficienty v lineární kombinaci jako na dvojrozměrnou funkci $F(k, l)$, $k = 0, 1, \dots, M-1$, $l = 0, 1, \dots, N-1$.

Příklad 1.2. Navrhněte bázi pro prostor obrazových signálů z příkladu 1.1 z předchozí podkapitoly.

Řešení. Je zřejmé, že hledaná báze musí obsahovat devět funkcí (obrazů) definovaných nad devíti pixely obrazu. Jednou z možností je zvolit funkce tak, že každá bude obsahovat vždy jednu hodnotu 1 a ostatní hodnoty budou 0. Každá z funkcí bude mít při tom hodnotu 1 na jiném místě. Bázi takto konstruovanou jistě velmi dobře znáte např. z práce s trojrozměrnými vektorovými prostory. Uvedené připomenutí se hodí. Na náš prostor obrazů majících 3×3 pixely můžeme totiž také nahlížet jako na prostor vektorů o devíti prvcích. Báze, kterou jsme nyní navrhli, je ortonormální. Později se seznámíme s bázemi, které jsou navrženy méně přímočaře. ▲

1.1.2 Skalární součin nad prostorem signálů

Řešení úloh se usnadní, jestliže nad prostorem signálů zavedeme *skalární součin*. Nechť \mathcal{S} je lineární prostor signálů. Skalární součin $\langle \varphi, \psi \rangle$ přiřazuje každým dvěma prvkům φ, ψ tohoto prostoru číslo (obecně se jedná o číslo komplexní) těchto vlastností:

$$\langle \varphi, \psi \rangle = \langle \psi, \varphi \rangle^* , \quad (1.6)$$

$$\langle a\varphi_1 + b\varphi_2, \psi \rangle = a\langle \varphi_1, \psi \rangle + b\langle \varphi_2, \psi \rangle , \quad (1.7)$$

$$\langle \varphi, \varphi \rangle \geq 0, \quad \langle \varphi, \varphi \rangle = 0 \Leftrightarrow \varphi = 0, \quad (1.8)$$

kde a, b jsou čísla (obecně opět komplexní) a $*$ označuje číslo komplexně sdružené. Prostor, v němž je zaveden skalární součin, se nazývá *unitárním prostorem*. Pomocí skalárního součinu lze zavést normu $\| \varphi \|$ prvku φ jako odmocninu ze skalárního součinu

$$\| \varphi \| = \sqrt{\langle \varphi, \varphi \rangle} . \quad (1.9)$$

Prvky φ, ψ unitárního prostoru nazveme ortogonálními, jestliže platí

$$\langle \varphi, \psi \rangle = 0 . \quad (1.10)$$

Nechť $\{\varphi_k\}$ je systém prvků unitárního prostoru. Tento systém nazveme ortogonálním, jestliže platí

$$\langle \varphi_k, \varphi_l \rangle \begin{cases} \neq 0, & k = l \\ = 0, & k \neq l \end{cases} . \quad (1.11)$$

Systém $\{\varphi_k\}$ nazveme ortonormálním, jestliže navíc platí

$$\langle \varphi_k, \varphi_l \rangle = \delta(k - l) , \quad (1.12)$$

kde $\delta(i)$ je tzv. *Kroneckerova funkce*, která je definována nad množinou celých čísel vztahem

$$\delta(i) = \begin{cases} 1, & i = 0 \\ 0, & i = \pm 1, \pm 2, \dots \end{cases} . \quad (1.13)$$

Uvažujme nyní unitární prostor \mathcal{S} , ve kterém je systém $\{\varphi_k\}$ bází. Jestliže je systém $\{\varphi_k\}$ ortogonální popř. ortonormální, pak hovoříme o ortogonální nebo o ortonormální bázi. K prověření úplnosti systému funkcí $\{\varphi_k\}$ v \mathcal{S} lze použít následující

větu. Systém funkcí $\{\varphi_k\}$ je úplný v \mathcal{S} právě tehdy, jestliže v \mathcal{S} neexistuje prvek (s výjimkou nulového), který by byl ortogonální ke všem prvkům systému $\{\varphi_k\}$ (větu uvádíme bez důkazu).

Zvolme nyní v prostoru \mathcal{S} diskretních signálů konkrétní skalární součin a ukažme důsledky této volby. Definičním oborem funkcí prostoru \mathcal{S} nechť je nejprve množina $\Omega = \{0, 1, \dots, M-1\}$ a $\varphi(m), \psi(m)$ nechť jsou funkce z \mathcal{S} . Skalární součin v prostoru \mathcal{S} zavedme takto

$$\langle \varphi(m), \psi(m) \rangle = \sum_{m=0}^{M-1} \varphi(m) \psi^*(m), \quad (1.14)$$

kde $*$ označuje komplexně sdruženou funkci. Na čtenáři ponecháváme, aby ukázal, že takto definovaný skalární součin má vlastnosti (1.6) až (1.8). Nechť $\{\varphi_k(m)\}$ je báze v prostoru \mathcal{S} . Podmínku (1.12) ortonormality lze nyní rozepsat ve tvaru

$$\sum_{m=0}^{M-1} \varphi_k(m) \varphi_l^*(m) = \delta(k-l). \quad (1.15)$$

Lze ukázat, že uvedenému vztahu je ekvivalentní vztah

$$\sum_{k=0}^{M-1} \varphi_k(m) \varphi_k^*(n) = \delta(m-n). \quad (1.16)$$

Analogicky lze postupovat také ve spojitém případě. Nechť \mathcal{S} je prostor spojitých signálů. V \mathcal{S} uvažujme bázi $\varphi(u, x)$, $x \in \Omega$, $u \in U$. Podmínka ortonormality báze má nyní tvar

$$\langle \varphi(u, x), \varphi(v, x) \rangle = \delta(u-v), \quad (1.17)$$

kde $\delta(u)$ je v tomto případě tzv. Diracův impuls. (I v dalším textu budeme podle potřeby symbolu δ používat k označení buď Kroneckerovy funkce nebo Diracova impulsu. O kterou funkci jde, bude vždy zřejmé z kontextu). Diracův impuls nabývá nenulové hodnoty pouze pro $u=0$, kde je jeho funkční hodnota nekonečno. Podrobněji o této funkci pojednáme v odstavci 1.2.3. Nechť $\varphi(x), \psi(x)$ jsou funkce z \mathcal{S} . V prostoru \mathcal{S} zavedeme skalární součin předpisem

$$\langle \varphi(x), \psi(x) \rangle = \int_{\Omega} \varphi(x) \psi^*(x) dx. \quad (1.18)$$

Ověření, že takto zavedený skalární součin má vlastnosti (1.6) až (1.8), opět ponecháváme čtenáři jako cvičení. Podmínku ortonormality báze můžeme nyní zapsat ve tvaru

$$\int_{\Omega} \varphi(u, x) \varphi^*(v, x) dx = \delta(u-v). \quad (1.19)$$

Opět lze ukázat, že uvedenému vztahu je ekvivalentní vztah

$$\int_U \varphi(u, x) \varphi^*(u, y) du = \delta(x - y). \quad (1.20)$$

Zobecnění dosud uvedeného na dvojrozměrné signály je přímočaré. V případě dvojrozměrných signálů máme $\Omega \subseteq E^2$. Vztahům (1.14) až (1.20) postupně odpovídají následující vztahy (vztahy (1.21), (1.24) podávají návod na výpočet skalárního součinu; vztahy (1.22), (1.23), (1.25), (1.26) vyjadřují podmínku ortonormality):

$$\langle \varphi(m, n), \psi(m, n) \rangle = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \varphi(m, n) \psi^*(m, n), \quad (1.21)$$

$$\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \varphi_{k_1, l_1}(m, n) \varphi_{k_2, l_2}^*(m, n) = \delta(k_1 - k_2) \delta(l_1 - l_2) = \delta(k_1 - k_2, l_1 - l_2), \quad (1.22)$$

$$\sum_{k=0}^{M-1} \sum_{l=0}^{N-1} \varphi_{k, l}(m_1, n_1) \varphi_{k, l}^*(m_2, n_2) = \delta(m_1 - m_2, n_1 - n_2), \quad (1.23)$$

$$\langle \varphi(x, y), \psi(x, y) \rangle = \iint_{\Omega} \varphi(x, y) \psi^*(x, y) dx dy, \quad (1.24)$$

$$\iint_{\Omega} \varphi(u_1, v_1, x, y) \varphi^*(u_2, v_2, x, y) dx dy = \delta(u_1 - u_2, v_1 - v_2), \quad (1.25)$$

$$\iint_U \varphi(u, v, x_1, y_1) \varphi^*(u, v, x_2, y_2) du dv = \delta(x_1 - x_2, y_1 - y_2). \quad (1.26)$$

Poznamenejme, že výše uvedených vztazích jsme použili dvojrozměrnou Kroneckerovu funkci, kterou lze definovat vztahem $\delta(i, j) = \delta(i) \delta(j)$, a dvojrozměrný Diracův impulz $\delta(x, y) = \delta(x) \delta(y)$.

1.1.3 Reprezentace signálu pomocí báze funkcí

V této podkapitole ukážeme, jak je možné vyjádřit obraz jako lineární kombinaci zvolených báze funkcí. Uvedený postup později vícekrát využijeme. Je základem transformací obrazu, např. transformace Fourierovy. Nechť \mathcal{S} je unitární prostor diskrétních signálů definovaných nad množinou M bodů, nechť $f(m)$ je signál z tohoto prostoru a $\{\varphi_k(m) | k = 0, 1, \dots, M-1\}$ nechť je ortonormální báze v \mathcal{S} . Ze vztahu (1.2) víme, že $f(m)$ je možné vyjádřit jako lineární kombinaci

$$f(m) = \sum_{l=0}^{M-1} F_l \varphi_l(m). \quad (1.27)$$

Provedeme-li skalární součin obou stran rovnice (1.27) s funkcí $\varphi_k(m)$ báze, pak získáme

$$\langle f(m), \varphi_k(m) \rangle = \sum_{l=0}^{M-1} F_l \langle \varphi_l(m), \varphi_k(m) \rangle. \quad (1.28)$$

S ohledem na předpokládanou ortonormalitu báze získáváme hledaný výsledek

$$F_k = \langle f(m), \varphi_k(m) \rangle. \quad (1.29)$$

Rozepíšeme-li skalární součin podle vztahu (1.14) a zapíšeme-li M -tici F_0, F_1, \dots, F_{M-1} formálně jako funkci, pak lze vztah (1.29) přepsat do tvaru

$$F(k) = \sum_{m=0}^{M-1} f(m) \varphi_k^*(m). \quad (1.30)$$

Vztah (1.30) podává návod, jak pro zadaný signál $f(m)$ a zadanou ortonormální bázi $\{\varphi_k(m)\}$ prostoru \mathcal{S} určit M -tici koeficientů $(F_0, F_1, \dots, F_{M-1})$, které signál reprezentují.

Analogicky je opět možné postupovat také pro spojité signály. Necht \mathcal{S} je unitární prostor jednorozměrných spojitých signálů, $\varphi(u, v)$ necht je ortonormální báze a $f(x)$ necht je signál v \mathcal{S} . Ze vztahu (1.3) víme, že f je možné vyjádřit jako lineární kombinaci

$$f(x) = \int_U F(v) \varphi(v, x) dv. \quad (1.31)$$

Provedeme-li skalární součin obou stran rovnice (1.31) s $\varphi(u, x)$, pak po úpravě získáme

$$F(u) = \langle f(x), \varphi(u, x) \rangle. \quad (1.32)$$

Rozepíšeme-li skalární součin podle (1.18), pak lze vztah (1.32) přepsat na

$$F(u) = \int_{\Omega} f(x) \varphi^*(u, x) dx. \quad (1.33)$$

Vztah (1.33) opět podává návod, jak pro zadaný signál $f(x)$ a pro zadanou ortonormální bázi $\varphi(u, x)$ určit funkci $F(u)$, která signál reprezentuje. Jak je již obvyklé, získáme podobně snadno také předpisy pro funkce $F(k, l)$, $F(u, v)$ použité ve vztazích (1.4) a (1.5):

$$F(k, l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \varphi_{k,l}^*(m, n), \quad (1.34)$$

$$F(u, v) = \iint_{\Omega} f(x, y) \varphi^*(u, v, x, y) dx dy. \quad (1.35)$$

Příklad 1.3. Promyslete, co by se stalo, kdyby báze, s níž pracujeme, nebyla ortonormální, ale nejprve pouze ortogonální a pak zcela obecná.

Řešení. Uvažujme diskrétní jednorozměrný případ. V případě ortogonální báze není situace o mnoho obtížnější než u báze ortonormální. Opět se snadno podaří nalézt explicitní vztah pro hledané koeficienty. Potíže ale nastanou v případě báze zcela obecné. Zde je pak nutné koeficienty určit řešením soustavy rovnic. Protože počet bázových funkcí je v praktických případech velký, je také velká soustava, která se má řešit. Teď je již jasné, proč v minulosti byly (a dodnes také jsou) ortonormální či ortogonální báze tak akcentovány. Řešení pro obecnou bázi dříve nebylo prakticky možné. ▲

1.2 Operace nad prostorem signálů

Matematickým modelem úprav obrazu jsou operace nad prostorem obrazových signálů. Průpravy získané v této podkapitole opět později bohatě využijeme. Necht \mathcal{S} je prostor signálů. Modelem úpravy signálu je unární nebo n -ární operátor \mathcal{O} realizující zobrazení $\mathcal{S} \rightarrow \mathcal{S}$, popřípadě $\mathcal{S}^n \rightarrow \mathcal{S}$. Jestliže, s ohledem na naši aplikaci, uvažujeme prostor dvojrozměrných signálů a jestliže $f(x, y), f_1(x, y), \dots, f_n(x, y) \in \mathcal{S}$ jsou prvky tohoto prostoru, pak můžeme úpravu signálu popsat vztahy

$$g(x, y) = \mathcal{O} \{f(x, y)\}, \quad (1.36)$$

$$g(x, y) = \mathcal{O} \{f_1(x, y), f_2(x, y), \dots, f_n(x, y)\}, \quad (1.37)$$

kde $g(x, y)$ je opět prvkem uvažovaného prostoru. Velmi často předpokládáme, že operátor \mathcal{O} má nějaké speciální vlastnosti; nejčastěji se jedná o linearitu a invarianci vůči posuvu. Pojďme o nich v následujících podkapitolách.

1.2.1 Lineární operace

Většina výsledků v teorii zpracování signálů je odvozena za předpokladu, že úprava signálu je popsána operátorem, který je lineární. Vysvětleme tento pojem. Necht $f(x, y), g(x, y) \in \mathcal{S}$ jsou signály; a, b necht jsou reálná nebo komplexní čísla. Operátor \mathcal{O} nazveme lineárním operátorem, jestliže platí

$$\mathcal{O} \{af(x, y) + bg(x, y)\} = a\mathcal{O} \{f(x, y)\} + b\mathcal{O} \{g(x, y)\}. \quad (1.38)$$

Uvedený vztah můžeme samozřejmě zobecnit pro libovolný počet sčítanců.

1.2.2 Operace invariantní vůči posuvu

Někdy požadujeme, aby měl operátor kromě linearitu ještě některé další vlastnosti. Velmi často to bývá invariance operátoru vůči posuvu, která je významnou vlastností

ve zpracování obrazu. Pokud bychom ji nepředpokládali, nebylo by možné, jak dále uvidíme, odvodit mnoho užitečných výsledků.

Nechť $f(x, y)$ je signál. Předpokládejme, že aplikací operátoru \mathcal{O} na signál f obdržíme signál g ; je tedy $g(x, y) = \mathcal{O}\{f(x, y)\}$. Operátor \mathcal{O} nazveme invariantním vůči posuvu, jestliže pro libovolný signál $f(x, y)$ a pro každou dvojici a, b vyplývá ze vztahu $g(x, y) = \mathcal{O}\{f(x, y)\}$, že platí také

$$g(x - a, y - b) = \mathcal{O}\{f(x - a, y - b)\} . \quad (1.39)$$

Výše uvedené můžeme vyslovit také méně formálně. Jestliže operátor invariantní vůči posuvu aplikujeme na dva vzájemně posunuté, ale jinak shodné signály, pak výsledkem operace jsou signály, které jsou také vzájemně posunuté, ale jinak shodné. Vzájemný posuv výsledných signálů je při tom stejný jako posuv signálů vstupních.

1.2.3 Diracův impulz

Mnoho pozoruhodných výsledků ve zpracování signálů bylo odvozeno za použití tzv. impulzové funkce - Diracova impulsu. Tento odstavec proto věnujeme uvedené funkci a jejím vlastnostem. Čtenáře, který se snad hlouběji zabývá o matematiku prosíme o určitou shovívavost k výkladu, který bude následovat. Výklad je zaměřen zejména na pochopení faktů, které jsou prakticky využitelné ve zpracování signálů. Pokud bychom požadovali výklad striktně matematicky korektní, bylo by zapotřebí vzít v úvahu, že Diracův impulz, přesně vzato, vlastně ani není funkcí, která by splňovala všechny požadavky, které na funkci obvykle klademe (nevyhovuje např. Lebesguově teorii integrování).

Pro další úvahy je užitečné nejprve zavést funkci $\text{rect}(x, y)$, která je definována předpisem

$$\text{rect}(x, y) = \begin{cases} 1, & (|x| \leq \frac{1}{2}) \cap (|y| \leq \frac{1}{2}) \\ 0, & \text{jinak} \end{cases} . \quad (1.40)$$

Dále zavedme funkci

$$\delta_n(x, y) = n^2 \text{rect}(nx, ny), \quad n = 1, 2, 3, \dots \quad (1.41)$$

tj.

$$\delta_n(x, y) = \begin{cases} n^2, & (|x| \leq \frac{1}{2n}) \cap (|y| \leq \frac{1}{2n}) \\ 0, & \text{jinak} \end{cases} .$$

S ohledem na to, jak byla funkce zavedena, je zřejmé, že platí

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \delta_n(x, y) dx dy = 1 . \quad (1.42)$$

Představme si nyní případ, kdy $n \rightarrow \infty$. Funkce $\delta_\infty(x, y)$ se nazývá Diracův impulz. Pro Diracův impulz budeme v dalším textu používat jednoduššího zápisu $\delta(x, y)$. Je

zřejmé, že pro všechna $(x, y) \neq (0, 0)$ je funkční hodnota Diracova impulsu $\delta(x, y) = 0$, v bodě $(0, 0)$ je hodnota nekonečno. Dále je zřejmé, že platí

$$\delta(-x, -y) = \delta(x, y), \quad (1.43)$$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \delta(x, y) dx dy = 1. \quad (1.44)$$

Pro odvození jedné z nejvýznačnějších vlastností Diracova impulsu mějme funkci $f(x, y)$ a uvažujme integrál (výraz na pravé straně následující rovnice snadno obdržíme s využitím dříve uvedené definice funkce δ_n)

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \delta_n(x, y) dx dy = n^2 \int_{-1/2n}^{1/2n} \int_{-1/2n}^{1/2n} f(x, y) dx dy. \quad (1.45)$$

Z výrazu na pravé straně rovnice (1.45) vidíme, že integrál na levé straně rovnice vyjadřuje střední hodnotu funkce $f(x, y)$ v oblasti nad čtvercem o stranách $1/n$ se středem v počátku souřadné soustavy. Pro $n \rightarrow \infty$ pak proto obdržíme střední hodnotu nad nekonečně malým čtvercem se středem v počátku, což je funkční hodnota $f(0, 0)$. Je tedy

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \delta(x, y) dx dy = f(0, 0). \quad (1.46)$$

Posuneme-li impuls do bodu (a, b) , pak analogicky máme

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \delta(x - a, y - b) dx dy = f(a, b). \quad (1.47)$$

Poznamenejme, že vlastnost popsaná vztahem (1.47) má zásadní význam a často se jí používá.

Konečně na závěr tohoto odstavce uvedeme ještě jednu vlastnost, kterou později upotřebíme. Vlastnost je vyjádřena vztahy

$$\begin{aligned} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp[-j2\pi(ux + vy)] du dv &= \delta(x, y), \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp[-j2\pi(ux + vy)] dx dy &= \delta(u, v), \end{aligned} \quad (1.48)$$

kde j je komplexní jednotka. Doporučujeme čtenáři, aby se alespoň neformálně pokusil uvedenou vlastnost zdůvodnit. Návod: Uvažte, že pro $x = y = 0$ se v prvním

případě jedná o integraci funkce nabývající hodnoty 1 nad nekonečnou oblastí. Pro $x \neq 0$, $y \neq 0$ se podle Eulerova vztahu $e^{j\varphi} = \cos \varphi + j \sin \varphi$ jedná o integraci funkcí typu $\cos \varphi \cos \psi$, $\cos \varphi \sin \psi$, $\sin \varphi \cos \psi$, $\sin \varphi \sin \psi$ ($\varphi = 2\pi ux$, $\psi = 2\pi vy$) nad nekonečnou oblastí. Obdobně lze postupovat i v případě druhého vztahu.

Příklad 1.4. Uvažujte Gaussovu funkci $\frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{x^2}{2\sigma^2})$, kterou asi dobře znáte z kurzů o pravděpodobnosti, kde popisuje tzv. normální rozložení hustoty pravděpodobnosti. Asi také víte, jaký význam má parametr σ a jak ovlivňuje tvar funkce. Pokuste se představit si limitní případ, kdy $\sigma \rightarrow 0$. Jakou funkci obdržíte?

Řešení. Održíte Diracův impulz $\delta(x)$. ▲

1.2.4 Lineární sumace bodových zdrojů, konvoluce

Již dříve jsme poznamenali, že na většinu úprav obrazu lze pohlížet jako na operátor. Zásadní otázkou zůstává, jak účinek operátoru vypočítat a jak operátor nějakým univerzálním a praktickým způsobem popsat. V této podkapitole ukážeme řešení obou problémů. Zavedeme k tomu pojem konvoluce, což je operace, která se ve zpracování obrazu používá mimořádně často. Jak se v průběhu výkladu ukáže, nelze konvoluce využít k výpočtu účinku zcela obecných operátorů, ale jen těch, které jsou lineární a invariantní vůči posunu. To však konvoluci na významu neubírá.

Uvažujme funkci $f(x, y)$. Podle vztahu (1.47) můžeme $f(x, y)$ s využitím Diracova impulzu zapsat ve tvaru

$$f(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(a, b) \delta(a - x, b - y) da db. \quad (1.49)$$

Protože se pro Diracův impulz používá také název *bodový zdroj*, říkáme, že jsme funkci $f(x, y)$ zapsali ve formě lineární sumace bodových zdrojů. Aby lépe vynikly vlastnosti uvedeného zápisu, zapíšeme integrál ze vztahu (1.49) na okamžik ve tvaru sumace. Dostaneme vztah

$$f(x, y) = \sum_i \sum_j f(a_i, b_j) \delta(a_i - x, b_j - y) \Delta a_i \Delta b_j. \quad (1.50)$$

Je užitečné poznamenat, že hodnota $f(a_i, b_j)$ ve vztahu (1.50) je na každém jednom intervalu konstantní, což později umožní výhodně využít předpokládané linearity operátoru.

Předpokládejme nyní, že na funkci $f(x, y)$ chceme aplikovat operátor \mathcal{O} . Budeme předpokládat, že operátor \mathcal{O} je lineární a invariantní vůči posuvu. Aplikací operátoru získáme výsledek $g(x, y)$. Máme

$$g(x, y) = \mathcal{O} \{f(x, y)\}. \quad (1.51)$$

Zapišeme-li funkci $f(x, y)$ podle vztahu (1.49) jako lineární sumu bodových zdrojů, dostaneme

$$g(x, y) = \mathcal{O}\{f(x, y)\} = \mathcal{O}\left\{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(a, b)\delta(a-x, b-y) da db\right\}. \quad (1.52)$$

Protože je podle předpokladu operátor \mathcal{O} lineární, dostáváme dále (viz též vztah (1.50))

$$g(x, y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(a, b)\mathcal{O}\{\delta(a-x, b-y) da db\} \quad (1.53)$$

a podle vztahu (1.43) také

$$g(x, y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(a, b)\mathcal{O}\{\delta(x-a, y-b) da db\}. \quad (1.54)$$

Předpokládejme dále, že podrobíme-li Diracův impulz $\delta(x, y)$ operaci \mathcal{O} , získáme funkci $h(x, y) = \mathcal{O}\{\delta(x, y)\}$. Funkce h se nazývá impulzová charakteristika operátoru \mathcal{O} (každý operátor invariantní vůči posuvu je touto charakteristikou zcela popsán). Protože je podle předpokladu operátor \mathcal{O} invariantní vůči posuvu, je $\mathcal{O}\{\delta(x-a, y-b)\} = h(x-a, y-b)$. S uvážením této skutečnosti je tedy

$$g(x, y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(a, b)h(x-a, y-b) da db. \quad (1.55)$$

Poznamenejme, že úpravou vztahu (1.53) na vztah (1.54) jsme převedli úlohu do tvaru, v němž je možné využít důsledků invariance operátoru \mathcal{O} vůči posuvu (předpoklad invariance totiž zajišťuje platnost vztahu $\mathcal{O}\{f(x-a, y-b)\} = g(x-a, y-b)$, obecně nikoli platnost vztahu $\mathcal{O}\{f(a-x, b-y)\} = g(a-x, b-y)$). Rovnice (1.55) má zásadní význam. Říká toto: Máme-li lineární operátor \mathcal{O} invariantní vůči posuvu, pak odezvu $\mathcal{O}\{f(x, y)\}$ na funkci $f(x, y)$ lze spočítat integrací na základě znalosti odezvy $h(x, y) = \mathcal{O}\{\delta(x, y)\}$ operátoru na Diracův impulz (tedy na základě jeho impulzové charakteristiky). Výraz na pravé straně rovnice (1.55) se nazývá konvolucí funkcí f a h . Operace konvoluce se označuje $*$. Zejména matematici rádi vidí, když je výsledná funkce vzniklá konvolucí zapsána ve tvaru $(f * h)(x, y)$. Lze zřejmě psát $(f * h)(x, y) = f(x, y) * h(x, y)$. Zatímco výraz na pravé straně výrazu říká, že máme dvě funkce, s nimiž teprve hodláme provést operaci konvoluce, výraz na straně levé odkazuje na již hotový výsledek. Z tohoto pohledu nepovažujeme rozdíl za příliš podstatný. Obě varianty však budeme, podle okolností, používat. Je tedy

$$g(x, y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(a, b)h(x-a, y-b) da db = (f * h)(x, y). \quad (1.56)$$

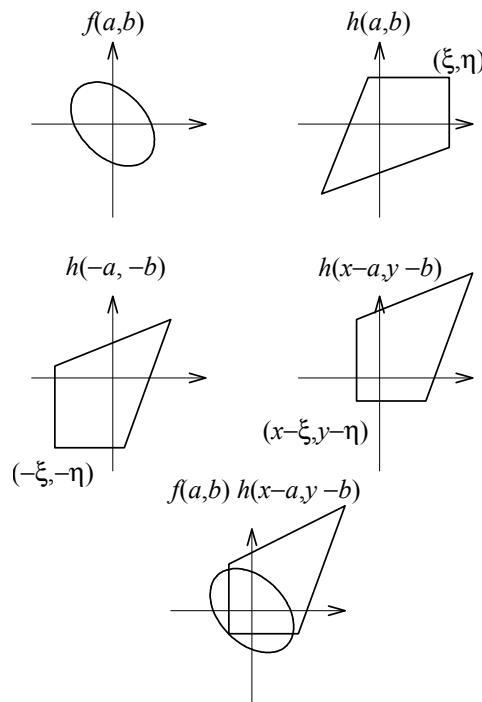
Po záměně proměnných $a' = x - a$, $b' = y - b$ a jejich opětovném přejmenování $a' \rightarrow a$, $b' \rightarrow b$ máme také

$$g(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(a, b) f(x - a, y - b) da db = (h * f)(x, y). \quad (1.57)$$

Je tedy zřejmé, že platí

$$(f * h)(x, y) = (h * f)(x, y) = f(x, y) * h(x, y) = h(x, y) * f(x, y). \quad (1.58)$$

Ze vztahu (1.56) vidíme, že konvoluce počítá výslednou hodnotu v bodě (x, y) jako „vážený průměr“ funkce f s váhovou funkcí h , která se často nazývá konvoluční maska nebo filtr (přesněji řečeno se jedná o impulzovou charakteristiku filtru; filtr je operátorem). Výpočet při tom probíhá tak, že se funkce h v rovině x, y otočí o 180° (důsledek členů se zápornými znaménky v (1.56)) a pak se posune o vektor (x, y) . Průběh výpočtu je ilustrován na obr. 1.1.



Obr. 1.1: K postupu výpočtu konvoluce

Pro úplnost poznamenejme, že vztahy pro jednorozměrnou konvoluci lze získat ze vztahů (1.56), (1.57) tak, že členy odpovídající jedné z dimenzí jednoduše vypustíme. Máme tedy

$$(f * h)(x) = (h * f)(x) = \int_{-\infty}^{+\infty} f(a)h(x - a) da = \int_{-\infty}^{+\infty} f(x - a)h(a) da. \quad (1.59)$$

1.2.5 Diskrétní konvoluce

Uvažujme nyní prostor diskretních obrazových signálů. Nechť je $\Omega = \{(m, n) \mid m = 0, 1, \dots, M-1; n = 0, 1, \dots, N-1\}$ a $f(m, n)$ a $h(m, n)$ nechť jsou signály z uvažovaného prostoru. Diskrétní konvoluci lze pak zapsat pomocí vztahů

$$(f * h)(m, n) = \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} f(r, s)h(m-r, n-s), \quad (1.60)$$

$$(h * f)(m, n) = \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} h(r, s)f(m-r, n-s). \quad (1.61)$$

Povšimněme si, že při výpočtu konvoluce budou pro funkce h popř. f ve vztazích (1.60), (1.61) vyžadovány také hodnoty vně rozsahu $0, 1, \dots, M-1; 0, 1, \dots, N-1$. Tento problém lze vyřešit dvěma způsoby. V obou doplníme definici funkcí f, h i mimo oblast Ω . Při tzv. lineární konvoluci předpokládáme, že funkce f, h mimo oblast Ω nabývají hodnoty 0. Při tzv. cyklické konvoluci funkce f, h naopak považujeme za periodické s periodou M, N . Výpočet cyklické konvoluce lze pak provést na základě vztahů

$$(f * h)(m, n) = \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} f(r, s)h((m-r)_M, (n-s)_N) \quad (1.62)$$

nebo

$$(h * f)(m, n) = \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} h(r, s)f((m-r)_M, (n-s)_N), \quad (1.63)$$

kde zápis $(p)_q$ znamená p modulo q . V praxi jsou užívány oba způsoby výpočtu (konvoluce lineární i cyklická). Volba závisí na tom, zda v daném okamžiku více vyhovuje představa, že signál nabývá mimo oblast Ω nulových hodnot nebo představa, že je signál periodický. V kapitole 2 ukážeme, že představa periodického signálu je užitečná zejména v souvislosti s Fourierovou transformací. Poznamenejme ještě, že kdyby bylo $m, n = -\infty, \dots, \infty$, pak by přinejmenším teoreticky k žádnému problému nedošlo. To může ilustrovat naše tvrzení z podkapitoly 1.1, že práce s nekonečně velkými obrazy může být někdy jednodušší než práce s obrazy konečné velikosti. Něco jiného je pak ale praktická implementace na počítači.

Předpokládejme, že je funkce h reprezentována maticí. Při praktických výpočtech často bývá funkce h nenulová pouze v relativně malém počtu bodů (je-li to možné, pak takto funkci h často záměrně volíme či zjednodušujeme, aby časová složitost výpočtu konvoluce byla přijatelná). Při použití dosud uvedených vztahů jsou nenulové hodnoty většinou bohužel soustředěny do rohů matice, což komplikuje algoritmizaci a snižuje efektivnost výpočtu. (To je proto, že $h(x, y)$ má být impulzovou charakteristikou operátoru, a tedy odezvou na $\delta(0, 0)$. Je-li bod $(0, 0)$ rohem obrazu, pak nezbyvá než předpokládat periodicitu; $h(x, y)$ pak nejspíš bude nenulové nejméně

v dalších třech exemplářích obrazu vyniklých periodicitou, a tedy i ve zbývajících třech rozích obrazu původního.) Z uvedených důvodů je proto užitečné členy v matici reprezentující funkci h přeskládat a odpovídajícím způsobem také modifikovat vztahy pro výpočet konvoluce. Předpokládejme pro jednoduchost, že hodnoty M , N jsou liché a položme $P = (M - 1)/2$, $Q = (N - 1)/2$. Zavedme funkci \tilde{h} , která je definována nad oblastí $\{(m, n) \mid m = -P, \dots, P; n = -Q, \dots, Q\}$, pomocí vztahu

$$\tilde{h}(m, n) = \begin{cases} h(-m, -n), & m \leq 0, n \leq 0 \\ h(M - m, -n), & m > 0, n \leq 0 \\ h(-m, N - n), & m \leq 0, n > 0 \\ h(M - m, N - n), & m > 0, n > 0 \end{cases} . \quad (1.64)$$

Funkce podle uvedeného předpisu jednoduše pouze získá užitečné hodnoty z rohů obrazu a dále provede středové převrácení tabulky získaných hodnot podle počátku (viz též obr. 1.1); ve vztahu pro konvoluci již proto později nebude zapotřebí znamének minus. S využitím funkce \tilde{h} lze pak pro výpočet konvoluce použít modifikovaného vztahu

$$(f * h)(m, n) = \sum_{r=-P}^P \sum_{s=-Q}^Q \tilde{h}(r, s) f(m + r, n + s) . \quad (1.65)$$

Předpokládejme, že funkční hodnoty $\tilde{h}(m, n)$ jsou nenulové pouze nad oblastí $\{(m, n) \mid m = -R, \dots, R; n = -S, \dots, S\}$, kde $R < P$, $S < Q$. Ve vztahu (1.65) tak proto stačí sčítat v mezích $-R \leq r \leq R$, $-S \leq s \leq S$ a k praktické reprezentaci funkce \tilde{h} na počítači stačí matice rozměru $(2R + 1) \times (2S + 1)$ (pro tuto matici se často používá termínu konvoluční maska).

Dále je zajímavé si povšimnout, že předpisem pro cyklickou konvoluci lze vypočítat též konvoluci lineární. Protože tato možnost má praktický význam, rozeberme ji podrobněji. Vyjděme ze vztahu (1.65), jímž se v praxi lineární konvoluce nejčastěji počítá. Uvažujme oblast $\tilde{\Omega} = \{(m, n) \mid m = 0, 1, \dots, M + R - 1; n = 0, 1, \dots, N + S - 1\}$ a rozšířme definici funkce f na celou oblast $\tilde{\Omega}$ tak, že funkci f položíme nad oblastí $\tilde{\Omega} \setminus \Omega$ rovnu nule. Nyní by již pro čtenáře mělo být snadným cvičením, aby ukázal, že lineární a cyklická konvoluce dají v tomto případě shodný výsledek.

Kapitola 2

Transformace obrazových signálů

Při zpracování signálů, a to včetně signálů obrazových, se často uplatňují techniky opírající se o transformace signálu. V této kapitole se budeme problematice transformací signálu věnovat podrobněji. Nejprve se zaměříme na obecný pohled na transformace. Dále se pak budeme věnovat transformaci Fourierově, která mezi transformacemi zaujímá významné postavení a v praxi se často uplatňuje. Konečně uvedeme i některé další transformace.

Transformací obecně rozumíme získání popisu signálu jako lineární kombinace funkcí jisté zvolené báze. Velmi často je vstupem transformace signál popsáný obrazovou funkcí přímo udávající intenzitu jasu v jednotlivých bodech obrazu. Nejprve ukážeme, že i tento nejběžnější způsob reprezentace obrazu lze interpretovat jako lineární kombinaci jistých bázových funkcí. Na transformaci lze pak tedy vždy pohlízet jako na proces, v němž jsme se z nějakých důvodů rozhodli změnit bázové funkce.

2.1 Báze δ

Uvažujme nejprve prostor \mathcal{S} diskrétních signálů, kde definičním oborem signálů je množina $\Omega = \{0, 1, \dots, M-1\}$. V tomto prostoru uvažujme bázi

$$\{\varphi_k(m) = \delta(k-m) \mid k, m = 0, 1, \dots, M-1\}. \quad (2.1)$$

Je zřejmé, že tato báze je ortonormální. V prostoru \mathcal{S} dále uvažujme signál popsáný funkcí $f(m)$ udávající intenzitu signálu v jednotlivých bodech. Ze vztahu (1.2) víme, že každý signál je možné vyjádřit jako lineární kombinaci bázových funkcí. Je tedy

$$f(m) = \sum_{k=0}^{M-1} F(k)\delta(k-m). \quad (2.2)$$

Hledejme nyní funkci $F(k)$. S využitím vztahu (1.30) máme

$$F(k) = \sum_{m=0}^{M-1} f(m)\delta(k-m) = f(k), \quad k = 0, 1, \dots, M-1. \quad (2.3)$$

Na základě vztahu (2.3) zjišťujeme, že pro bázi dle (2.1) platí $F(k) = f(k)$. Tento závěr není překvapivý (je spíše triviální). Chtěli jsme zde pouze ukázat, že i běžný způsob praktické reprezentace diskrétního signálu pomocí M -tice $(f_0, f_1, \dots, f_{M-1})$ intenzit lze interpretovat z pohledu použití bazových funkcí.

Analogicky můžeme postupovat také v prostoru spojitých signálů. Zvolme bázi ve tvaru

$$\varphi(u, x) = \delta(u - x). \quad (2.4)$$

Opět je zřejmé, že takto zvolená báze je ortonormální. Uvažujme nyní v prostoru \mathcal{S} spojitý signál $f(x)$ definovaný nad oblastí Ω . Ze vztahu (1.3) víme, že tento signál je možné zapsat ve tvaru lineární kombinace

$$f(x) = \int_U F(u) \delta(u - x) du. \quad (2.5)$$

Na základě vztahu (1.33) pak máme

$$F(u) = \int_{\Omega} f(x) \delta(u - x) dx = f(u). \quad (2.6)$$

Vztah (2.6) ukazuje, že při volbě báze dle (2.4) je hledaná funkce F rovna přímo funkci f . I zde jsme chtěli pouze ukázat, že i reprezentace signálu funkcí f má interpretaci z pohledu bazových funkcí. Pro bázi dle vztahu (2.1), (2.4) budeme v dalším textu používat názvu báze δ .

2.2 Transformace signálu

Mějme unitární prostor signálů. V tomto prostoru zvolme bázi. Transformací rozumíme pochod, kdy pro signál popsany pomocí obrazové funkce (tedy jako lineární kombinace funkcí báze δ , jak jsme ukázali v předchozím odstavci) určujeme jeho popis pomocí zvolené báze (určujeme tedy koeficienty do lineární kombinace zvolených bazových funkcí). Zpětnou (inverzní) transformací pak nazýváme pochod opačný, kdy pro signál popsany pomocí zvolené báze určujeme obrazovou funkci.

Uvažujme nejprve diskrétní jednorozměrné signály. Transformací signálu $f(m)$ rozumíme nalezení funkce $F(k)$ ze vztahu (1.2), která signál reprezentuje s ohledem na zvolenou bázi $\{\varphi_k\}$. Je-li báze $\{\varphi_k\}$ ortonormální (což je výhodné a časté), pak lze pro výpočet koeficientů použít vztahu (1.30). Známe-li naopak funkci $F(k)$, pak lze obrazovou funkci $f(m)$ určit jako lineární kombinaci (1.2). Pro diskrétní dvojrozměrné signály je odpovídající zobecnění popsáno vztahy (1.34), (1.4).

Transformaci diskrétního dvojrozměrného signálu nazveme separabilní, jestliže pro bazové funkce platí

$$\varphi_{k,l}(m, n) = \varphi_k(m) \varphi_l(n). \quad (2.7)$$

Je-li transformace separabilní, pak ze vztahů (1.34) a (2.7) plyne

$$F(k, l) = \sum_{m=0}^{M-1} \left[\sum_{n=0}^{N-1} f(m, n) \varphi_l^*(n) \right] \varphi_k^*(m). \quad (2.8)$$

Vztah (2.8) ukazuje, že jestliže je transformace separabilní, pak ji lze provádět odděleně po řádcích a po sloupcích bodů oblasti Ω , nad níž je signál definován. To je prakticky významné, protože lze takto, podle okolností, zkrátit čas výpočtu.

Předpokládejme nyní, že diskrétní dvojrozměrný signál je reprezentován jednorozměrným vektorem \mathbf{f} (stačí zvolit pravidlo, podle kterého hodnoty $f(m, n)$ umísťujeme do jednorozměrného vektoru; obraz lze např. procházet po řádcích) a podobně předpokládejme, že i výsledek transformace je reprezentován jednorozměrným vektorem \mathbf{F} . Každý z těchto vektorů má MN prvků. Ze vztahu (1.34) vidíme, že každý prvek vektoru \mathbf{F} je lineární kombinací prvků z \mathbf{f} . Lze tedy transformaci (1.34) vyjádřit též maticovým zápisem

$$\mathbf{F} = \mathbf{\Phi} \mathbf{f}, \quad (2.9)$$

$$\mathbf{f} = \mathbf{\Phi}^{-1} \mathbf{F}, \quad (2.10)$$

kde matice $\mathbf{\Phi}$ je rozměru $MN \times MN$. Poznamenejme, že je-li báze $\varphi_{k,l}(m, n)$ ortonormální, pak platí

$$\mathbf{\Phi}^{-1} = \mathbf{\Phi}^{*\top}. \quad (2.11)$$

Jestliže je transformace separabilní (vztah (2.7)), pak lze nalézt maticový zápis, který je co do rozměru matic poněkud úspornější. Necht' je tentokrát signál před a po transformaci reprezentován maticemi \mathbf{f} a \mathbf{F} , které jsou rozměru $M \times N$. Separabilní transformaci lze pak vyjádřit ve tvaru (praktický příklad uvedeme později v souvislosti s výpočtem Fourierovy transformace)

$$\mathbf{F} = \mathbf{\Phi}_C \mathbf{f} \mathbf{\Phi}_R^\top, \quad (2.12)$$

$$\mathbf{f} = \mathbf{\Phi}_C^{-1} \mathbf{F} \mathbf{\Phi}_R^{\top-1}. \quad (2.13)$$

Podle očekávání lze analogicky postupovat také v prostoru spojitých signálů. Necht' \mathcal{S} je takový prostor. V \mathcal{S} zvolme bázi $\varphi(u, x)$. Dále v \mathcal{S} uvažujme signál $f(x)$. Transformací signálu rozumíme stanovení funkce $F(u)$, která uvažovaný signál reprezentuje s ohledem na zvolenou bázi. Splňuje-li báze podmínku ortonormality (1.19), pak lze pro výpočet této funkce použít vztahu (1.33). Známe-li naopak funkci $F(u)$, lze podle vztahu (1.3) zapsat funkci $f(x)$ jako lineární kombinaci. Pro spojitě dvojrozměrné signály lze obdobně uplatnit vztahy (1.35), (1.5).

2.3 Fourierova transformace

Ústřední místo mezi transformacemi signálu zaujímá transformace Fourierova. V této podkapitole se zaměříme nejprve na Fourierovu transformaci spojitou, pak na transformaci diskrétní.

2.3.1 Spojitá Fourierova transformace

Uvažujme prostor signálů, kde oblast Ω definice signálu je dvojrozměrný euklidovský prostor E^2 . Fourierovu transformaci obdržíme, jestliže bázi volíme ve tvaru

$$\varphi(u, v, x, y) = \exp [j2\pi(ux + vy)] . \quad (2.14)$$

Ponecháváme na čtenáři, aby si sám ověřil, že uvedená báze je ortonormální - tj., že ve shodě se vztahem (1.25) platí

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp [j2\pi(u_1x + v_1y)] \exp [-j2\pi(u_2x + v_2y)] dx dy = \delta(u_1 - u_2, v_1 - v_2) . \quad (2.15)$$

(Návod: Použijte vlastnost (1.48) Diracova impulzu.) Na základě dříve odvozeného vztahu (1.35), (1.5) můžeme tedy pro spojitou Fourierovu transformaci dvojrozměrných signálů psát

$$\begin{aligned} F(u, v) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \exp [-j2\pi(ux + vy)] dx dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) [\cos 2\pi(ux + vy) - j \sin 2\pi(ux + vy)] dx dy . \end{aligned} \quad (2.16)$$

$$\begin{aligned} f(x, y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(u, v) \exp [j2\pi(ux + vy)] du dv \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(u, v) [\cos 2\pi(ux + vy) + j \sin 2\pi(ux + vy)] du dv . \end{aligned} \quad (2.17)$$

Ze vztahu (2.16) vyplývá, že obecně je $F(u, v)$ komplexní funkce. Rozepsáním na reálnou a imaginární složku máme

$$F(u, v) = R(u, v) + jI(u, v) . \quad (2.18)$$

Pro amplitudu $|F(u, v)|$ a fázový posuv $\Phi(u, v)$ pak platí

$$\begin{aligned} |F(u, v)| &= \sqrt{R^2(u, v) + I^2(u, v)} , \\ \Phi(u, v) &= \arctan \left(\frac{I(u, v)}{R(u, v)} \right) . \end{aligned} \quad (2.19)$$

Hodnota $|F(u, v)|$ vyjadřuje, jak je frekvence u, v obsažena v původní funkci $f(x, y)$. Funkce $|F(u, v)|^2$ se nazývá *energetické spektrum signálu* $f(x, y)$. Vztahy (2.16),

(2.17) provádí zobrazení komplexního signálového prostoru na sebe a jsou tedy operátorem nad tímto prostorem. Říkáme, že funkce $F(u, v)$ je Fourierovou transformací funkce $f(x, y)$ a funkce $f(x, y)$ je zpětnou (inverzní) Fourierovou transformací funkce $F(u, v)$. Tyto skutečnosti stručně zapisujeme takto

$$F(u, v) = \mathcal{F}\{f(x, y)\}, \quad f(x, y) = \mathcal{F}^{-1}\{F(u, v)\}. \quad (2.20)$$

Poznamenejme ještě, že při zpracování obrazového signálu $f(x, y)$ chápeme frekvenci jako „frekvenci délkovou“, tj. kolik period se vejde na jednotku délky obrazu.

Je dále zajímavé poukázat na souvislost Fourierovy transformace s Fourierovými řadami, které jsou dobře známé ze základního kurzu matematiky. Nechť $f(x, y)$ je periodická funkce s periodou l_x ve směru osy x a periodou l_y ve směru osy y . Fourierovou řadou funkce $f(x, y)$ pak nazýváme trigonometrickou řadu

$$\begin{aligned} f(x, y) &= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{mn} \exp \left[+j2\pi \left(\frac{mx}{l_x} + \frac{ny}{l_y} \right) \right] \\ &= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{mn} \left[\cos 2\pi \left(\frac{mx}{l_x} + \frac{ny}{l_y} \right) + j \sin 2\pi \left(\frac{mx}{l_x} + \frac{ny}{l_y} \right) \right], \end{aligned} \quad (2.21)$$

kde

$$c_{mn} = \frac{1}{l_x l_y} \int_{-l_x/2}^{l_x/2} \int_{-l_y/2}^{l_y/2} f(x, y) \exp \left[-j2\pi \left(\frac{mx}{l_x} + \frac{ny}{l_y} \right) \right] dx dy. \quad (2.22)$$

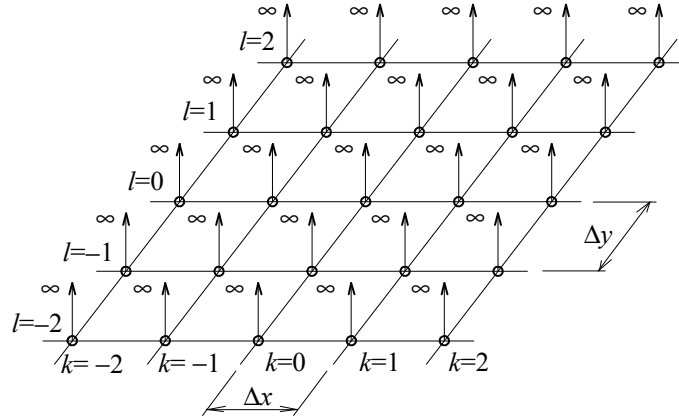
K přechodu od Fourierovy řady k Fourierově transformaci stačí obecnou (a tedy neperiodickou) funkci považovat za speciální případ funkce periodické, kdy je perioda nekonečně dlouhá, a dále vyjít z předpisu (2.21) pro Fourierovu řadu. Uvažujme periodickou funkci $f(x, y)$ s periodami l_x, l_y . Jestliže se periody l_x, l_y prodlužují, pak se podle vztahu (2.21) jednotlivé složky spektra navzájem přibližují, až je nakonec při nekonečně dlouhých periodách vzdálenost mezi sousedními složkami spektra nekonečně malá. Diskrétní spektrum se tak změnilo ve spektrum spojité.

Pro ilustraci spojité Fourierovy transformace uvedeme několik příkladů, jejichž výsledků později použijeme.

Příklad 2.1. Najdeme Fourierovu řadu pro funkci $f(x, y)$, která je tvořena nekonečným polem Diracových impulzů, které se opakují ve vzdálenostech Δ_x, Δ_y . (Výsledek řešení tohoto příkladu budeme později potřebovat v kapitole o vzorkování.)

Řešení. Funkce, kterou máme transformovat, má tvar

$$f(x, y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta(x - k\Delta_x, y - l\Delta_y).$$



Obr. 2.1 Pole Diracových impulzů

Funkce je vyobrazena na obr. 2.1. Podle vztahu (2.22) máme

$$\begin{aligned}
 c_{mn} &= \frac{1}{\Delta_x \Delta_y} \int_{-\Delta_x/2}^{\Delta_x/2} \int_{-\Delta_y/2}^{\Delta_y/2} \delta(x - k\Delta_x, y - l\Delta_y) \exp \left[-j2\pi \left(\frac{mx}{\Delta_x} + \frac{ny}{\Delta_y} \right) \right] dx dy \\
 &= \frac{1}{\Delta_x \Delta_y} .
 \end{aligned}$$

Vztah pro c_{mn} jsme obdrželi s uvážením skutečnosti, že integrujeme po délce jedné periody; v našem případě tedy od $-\Delta_x/2$ do $+\Delta_x/2$ a od $-\Delta_y/2$ do $+\Delta_y/2$, kde $\delta(x - k\Delta_x, y - l\Delta_y)$ nabývá nenulové hodnoty pouze pro $k = l = 0$. Dále jsme využili vlastnosti (1.46) Diracova impulzu. Pro hledanou řadu pak podle vztahu (2.21) vychází

$$f(x, y) = \frac{1}{\Delta_x \Delta_y} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \exp \left[j2\pi \left(\frac{mx}{\Delta_x} + \frac{ny}{\Delta_y} \right) \right] .$$

▲

Příklad 2.2. Nalezneme Fourierovu transformaci funkce $\delta_n(x, y)$, kterou jsme zavedli v podkapitole 1.2.3. Výsledek řešení příkladu je pro zpracování signálu významný a mnohokrát jej později využijeme.

Řešení. Na základě předpisu (2.16) postupně dostáváme

$$\begin{aligned} F(u, v) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta_n(x, y) \exp[-j2\pi(ux + vy)] dx dy \\ &= n^2 \int_{-1/2n}^{1/2n} \int_{-1/2n}^{1/2n} \exp[-j2\pi(ux + vy)] dx dy \\ &= \left[n \int_{-1/2n}^{1/2n} \exp[-j2\pi ux] dx \right] \left[n \int_{-1/2n}^{1/2n} \exp[-j2\pi vy] dy \right]. \end{aligned}$$

S využitím Eulerova vztahu $e^{j\varphi} = \cos \varphi + j \sin \varphi$ a po integraci (oba integrály v hranatých závorkách mají stejný tvar) pak vyjde

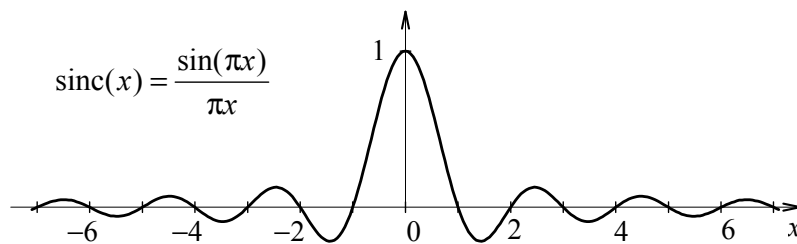
$$F(u, v) = n^2 \frac{\sin(\pi u/n)}{\pi u} \frac{\sin(\pi v/n)}{\pi v}.$$

Zavedeme dále funkci (obr. 2.2)

$$\text{sinc}(u, v) = \text{sinc}(u)\text{sinc}(v) = \frac{\sin(\pi u)}{\pi u} \frac{\sin(\pi v)}{\pi v}.$$

Pak vychází

$$F(u, v) = \mathcal{F}\{\delta_n(x, y)\} = \frac{\sin(\frac{\pi u}{n})}{\frac{\pi u}{n}} \frac{\sin(\frac{\pi v}{n})}{\frac{\pi v}{n}} = \text{sinc}\left(\frac{u}{n}, \frac{v}{n}\right). \quad (2.23)$$



Obr. 2.2 Funkce $\text{sinc}(x)$

Konstatujeme, že obdélníkový impulz $\delta_n(x, y)$ má nekonečné spojité reálné frekvenční spektrum. Amplituda složek spektra je vyjádřena funkcí $\text{sinc}(u/n, v/n)$. Protože se s funkcí sinc budeme setkávat velmi často, uvádíme na obr. 2.2 graf funkce $\text{sinc}(x)$. ▲

Příklad 2.3. Nalezneme Fourierovu transformaci Diracova impulzu $\delta(x, y)$.

Řešení. Využijeme výsledku z předchozího příkladu, kde položíme $n \rightarrow \infty$. Máme proto

$$\mathcal{F}\{\delta(x, y)\} = \lim_{n \rightarrow \infty} \left[\operatorname{sinc} \left(\frac{u}{n}, \frac{v}{n} \right) \right] = 1.$$

Frekvenční spektrum Diracova impulzu $\delta(x, y)$ je opět spojité reálné a nekonečné. Amplituda všech složek spektra je konstantní a je rovna 1. ▲

Příklad 2.4. Nalezneme Fourierovu transformaci funkce $f(x, y)$, která je tvořena nekonečným polem Diracových impulzů opakujících se ve vzdálenostech Δ_x, Δ_y .

Řešení. Funkci lze zapsat ve tvaru

$$f(x, y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta(x - k\Delta_x, y - l\Delta_y).$$

Uvedená funkce je vyobrazena na obr. 2.1. V příkladě 2.1 jsme pro tuto funkci již našli Fourierovu řadu (funkce je periodická). Abychom obdrželi „pěkný“ výsledek, provedeme nyní místo Fourierovy transformace původní funkce transformaci jejího Fourierova rozvoje, který jsme obdrželi v příkladu 2.1. Z příkladu 2.1 máme

$$f(x, y) = \frac{1}{\Delta_x \Delta_y} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \exp \left[j2\pi \left(\frac{mx}{\Delta_x} + \frac{ny}{\Delta_y} \right) \right].$$

Dosazením uvedeného vztahu do předpisu (2.16) máme

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\Delta_x \Delta_y} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \exp \left[j2\pi \left(\frac{mx}{\Delta_x} + \frac{ny}{\Delta_y} \right) \right] \exp [-j2\pi(ux + vy)] dx dy.$$

Po záměně pořadí integrování a sumace a po úpravě dostaneme

$$F(u, v) = \frac{1}{\Delta_x \Delta_y} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp \left\{ -j2\pi \left[\left(u - \frac{m}{\Delta_x} \right) x + \left(v - \frac{n}{\Delta_y} \right) y \right] \right\} dx dy.$$

S uvážením vlastnosti (1.48) Diracova impulzu pak konečně obdržíme vztah

$$F(u, v) = \frac{1}{\Delta_x \Delta_y} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta \left(u - \frac{m}{\Delta_x}, v - \frac{n}{\Delta_y} \right).$$

Vidíme, že Fourierovým obrazem nekonečného pole Diracových impulzů vzdálených vzájemně o Δ_x, Δ_y je opět nekonečné pole Diracových impulzů, kde jsou impulzy vzájemně vzdáleny o $1/\Delta_x, 1/\Delta_y$. ▲

Příklad 2.5. Ukažte, že Fourierovou transformací funkce $\cos(2\pi ax)$ je funkce $[\delta(u - a) + \delta(u + a)]/2$.

Řešení. Příklad je zajímavý, protože ilustruje, co Fourierova transformace počítá. Promyslete nejprve, jaký výsledek byste na základě toho, co o Fourierově transformaci víte, intuitivně očekávali. Provedení příkladu pak ponecháváme čtenáři jako cvičení. Návod: Použijte vztahu $\cos \varphi = (e^{j\varphi} + e^{-j\varphi})/2$ a dále použijte vlastnosti (1.48) Diracova impulzu. Z příkladu vidíme, že spektrum funkce $\cos(2\pi ax)$ obsahuje pouze složky $u = a$, $u = -a$. To je současně výsledek, který jste předem pravděpodobně očekávali. ▲

2.3.2 Vlastnosti spojitě Fourierovy transformace

V tomto odstavci uvedeme několik vlastností Fourierovy transformace, které se ve zpracování obrazu často uplatňují. Protože důkazy uvedených vlastností jsou vesměs jednoduché, ponecháváme je čtenáři jako cvičení (zpravidla však uvádíme stručný návod, jak při důkazu postupovat). Ve shodě s dříve používaným značením je i zde $F(u, v) = \mathcal{F}\{f(x, y)\}$ Fourierův obraz funkce $f(x, y)$, $G(u, v) = \mathcal{F}\{g(x, y)\}$ Fourierův obraz funkce $g(x, y)$; a, b jsou čísla. Obecně mohou být funkce $f(x, y)$, $g(x, y)$ i čísla a, b komplexní.

a) *Linearita:* Fourierova transformace je lineární, protože platí

$$\begin{aligned} \mathcal{F}\{af(x, y) + bg(x, y)\} &= a\mathcal{F}\{f(x, y)\} + b\mathcal{F}\{g(x, y)\} \\ &= aF(u, v) + bG(u, v). \end{aligned} \quad (2.24)$$

K důkazu této vlastnosti použijte předpis (2.16) pro Fourierovu transformaci. Dále již vlastnost přímo vyplývá z vlastností integrování.

b) *Měřítko:* Necht α, β jsou reálná čísla. Pak platí

$$\mathcal{F}\{f(\alpha x, \beta y)\} = F(u/\alpha, v/\beta)/|\alpha\beta|. \quad (2.25)$$

K důkazu této vlastnosti vyjděte z předpisu pro Fourierovu transformaci a zaveďte transformaci proměnných $x' = \alpha x$, $y' = \beta y$.

c) *Posuv:*

$$\mathcal{F}\{f(x - x_o, y - y_o)\} = F(u, v) \exp[-j2\pi(ux_o + vy_o)]. \quad (2.26)$$

K důkazu této vlastnosti vyjděte z předpisu pro Fourierovu transformaci a zaveďte transformaci proměnných $x' = x - x_o$, $y' = y - y_o$. Podobnou vlastnost lze nalézt také ve frekvenční doméně. Má tvar

$$\mathcal{F}^{-1}\{F(u - u_o, v - v_o)\} = f(x, y) \exp[j2\pi(u_o x + v_o y)]. \quad (2.27)$$

Vlastnost (2.27) bývá nazývána modulací. Ponecháváme na čtenáři, aby promyslel oprávněnost tohoto názvu.

d) *Reflexe*: Platí následující vztahy

$$\begin{aligned}\mathcal{F}\{f(-x, y)\} &= F(-u, v), & \mathcal{F}\{f(x, -y)\} &= F(u, -v), \\ \mathcal{F}\{f(-x, -y)\} &= F(-u, -v)\end{aligned}\quad (2.28)$$

a také

$$\mathcal{F}\{f^*(-x, -y)\} = F^*(u, v), \quad (2.29)$$

kde * označuje komplexně sdruženou veličinu. K důkazu vlastností (2.28) opět postačí substituce $x' = -x$, $y' = -y$. K důkazu vlastnosti (2.29) použijte předpisu (2.16) pro $F(u, v)$, proveďte komplexní sdružení a výsledek dále upravte.

e) *Konvoluční teorém*:

$$\mathcal{F}\{f(x, y) * g(x, y)\} = F(u, v)G(u, v), \quad (2.30)$$

$$\mathcal{F}\{f(x, y)g(x, y)\} = F(u, v) * G(u, v). \quad (2.31)$$

Vlastnost (2.30), (2.31) je významná a později ji mnohokrát využijeme. Říká, že Fourierovým obrazem konvoluce funkcí f, g je součin jejich Fourierových obrazů F, G . Podobně je Fourierovým obrazem součinu funkcí f, g konvoluce Fourierových obrazů F, G obou funkcí. S ohledem na mimořádný význam uvedených tvrzení provedeme tentokrát i důkaz.

Důkaz. Dokážeme vlastnost (2.30). Důkaz druhé vlastnosti lze provést obdobně. Víme, že konvoluce funkcí je definována předpisem

$$f(x, y) * g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b)g(x - a, y - b) da db.$$

Pro Fourierovu transformaci konvoluce pak podle předpisu (2.16) máme

$$\begin{aligned}\mathcal{F}\{f(x, y) * g(x, y)\} &= \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b)g(x - a, y - b) da db \right] \exp[-j2\pi(ux + vy)] dx dy.\end{aligned}$$

Záměnou pořadí integrace dostáváme

$$\begin{aligned}\mathcal{F}\{f(x, y) * g(x, y)\} &= \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b) \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x - a, y - b) \exp[-j2\pi(ux + vy)] dx dy \right] da db.\end{aligned}$$

Uplatněním vlastnosti (2.26) získáme

$$\mathcal{F}\{f(x, y) * g(x, y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b)G(u, v) \exp[-j2\pi(ua + vb)] da db.$$

Tento vztah lze pak již konečně upravit na tvar

$$\begin{aligned}\mathcal{F}\{f(x, y) * g(x, y)\} &= G(u, v) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b) \exp[-j2\pi(ua + vb)] da db \\ &= F(u, v)G(u, v).\end{aligned}$$

f) *Parsevalův teorém*: Předpokládejme, že obecně jsou funkce $f(x, y)$, $g(x, y)$ komplexní a použijme symbolu $*$ pro označení funkce komplexně sdružené. Parsevalův teorém tvrdí následující

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)g^*(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v)G^*(u, v) du dv. \quad (2.32)$$

Ve speciálním případě, kdy je $f(x, y) = g(x, y)$, přechází uvedený výraz na tvar

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x, y)|^2 dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |F(u, v)|^2 du dv. \quad (2.33)$$

Výraz na levé straně vztahu (2.33) může být interpretován jako celkový výkon signálu $f(x, y)$. Výraz $|F(u, v)|^2$ má proto význam výkonového spektra signálu $f(x, y)$.

g) *Korelační teorém*: Opět předpokládejme, že obecně jsou funkce $f(x, y)$, $g(x, y)$ komplexní. Křížová korelace R_{fg} a autokorelace R_{ff} deterministických funkcí $f(x, y)$ a $g(x, y)$ jsou definovány pomocí vztahů

$$R_{fg}(a, b) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^*(x - a, y - b)g(x, y) dx dy, \quad (2.34)$$

$$\begin{aligned}R_{ff}(a, b) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^*(x - a, y - b)f(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x + a, y + b)f^*(x, y) dx dy.\end{aligned} \quad (2.35)$$

Ne náhodou uvedené vztahy připomínají čtenáři vztah pro konvoluci. Jediný rozdíl je skutečně ve znaménku členu $(x - a, y - b)$. Také tvrzení korelačního teorému koresponduje s tvrzením teorému konvolučního. Korelační teorém totiž tvrdí:

$$\mathcal{F}\{R_{fg}\} = F^*(u, v)G(u, v), \quad (2.36)$$

$$\mathcal{F}\{R_{ff}\} = F^*(u, v)F(u, v) = |F(u, v)|^2. \quad (2.37)$$

Důkaz vztahu (2.36) lze provést podobně jako u konvolučního teorému. Vztah (2.37) vyplývá ze vztahu (2.36).

h) Derivace: Pro Fourierův obraz derivace funkce platí vztahy

$$\mathcal{F}\left\{\frac{\partial f(x, y)}{\partial x}\right\} = j2\pi u F(u, v), \quad \mathcal{F}\left\{\frac{\partial f(x, y)}{\partial y}\right\} = j2\pi v F(u, v). \quad (2.38)$$

K důkazu této vlastnosti stačí vyjít z definice Fourierovy transformace a použít integrace per partes. Při tom předpokládáme, že platí $f(-\infty, y) = f(\infty, y) = f(x, -\infty) = f(x, \infty) = 0$. Analogicky lze také odvodit např. vztah pro Fourierův obraz laplaciánu funkce. Dostaneme

$$\mathcal{F}\left\{\frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2}\right\} = -4\pi^2(u^2 + v^2)F(u, v). \quad (2.39)$$

2.3.3 Diskrétní Fourierova transformace

Uvažujme prostor signálů, kde je oblast Ω , nad níž jsou signály definovány, tvořena množinou bodů, které indexujeme pomocí dvojice indexů $m = 0, 1, \dots, M-1$, $n = 0, 1, \dots, N-1$. Diskrétní Fourierovu transformaci v tomto prostoru obdržíme, jestliže bázi volíme ve tvaru

$$\begin{aligned} \varphi_{k,l}(m, n) &= \frac{1}{\sqrt{MN}} \exp\left[j2\pi\left(\frac{mk}{M} + \frac{nl}{N}\right)\right] \\ k &= 0, 1, \dots, M-1, \\ l &= 0, 1, \dots, N-1. \end{aligned} \quad (2.40)$$

Opět ponecháváme na čtenáři, aby ověřil, že takto zvolená báze je ortonormální - tj., že platí

$$\begin{aligned} \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \exp\left[j2\pi\left(\frac{mk_1}{M} + \frac{nl_1}{N}\right)\right] \exp\left[-j2\pi\left(\frac{mk_2}{M} + \frac{nl_2}{N}\right)\right] \\ = \delta(k_1 - k_2, l_1 - l_2). \end{aligned} \quad (2.41)$$

Podle vztahu (1.34) pak pro diskrétní Fourierovu transformaci platí

$$\begin{aligned} F(k, l) &= \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \exp\left[-j2\pi\left(\frac{mk}{M} + \frac{nl}{N}\right)\right], \\ k &= 0, 1, \dots, M-1, \\ l &= 0, 1, \dots, N-1. \end{aligned}$$

V souladu se vztahem (1.4) pro zpětnou Fourierovu transformaci dostáváme

$$f(m, n) = \frac{1}{\sqrt{MN}} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} F(k, l) \exp \left[j2\pi \left(\frac{mk}{M} + \frac{nl}{N} \right) \right]. \quad (2.42)$$

Předpokládejme nyní, že signál je reprezentován maticí \mathbf{f} a jeho Fourierův obraz maticí \mathbf{F} . Obě matice mají rozměr $M \times N$. Protože Fourierova transformace je separabilní, lze ji zapsat podle vztahu (2.12) ve tvaru

$$\mathbf{F} = \Phi_{\mathbf{C}} \mathbf{f} \Phi_{\mathbf{R}}. \quad (2.43)$$

Zavedme označení

$$W_M = \exp \left[\frac{-j2\pi}{M} \right]. \quad (2.44)$$

Pro matice $\Phi_{\mathbf{R}}$, $\Phi_{\mathbf{C}}$ pak platí

$$\Phi_{\mathbf{R}} = \frac{1}{\sqrt{N}} \begin{bmatrix} W_N^0 & W_N^0 & W_N^0 & \cdots & W_N^0 \\ W_N^0 & W_N^1 & W_N^2 & \cdots & W_N^{N-1} \\ W_N^0 & W_N^2 & W_N^4 & \cdots & W_N^{2(N-1)} \\ \cdots & \cdots & \cdots & \ddots & \cdots \\ W_N^0 & W_N^{N-1} & W_N^{2(N-1)} & \cdots & W_N^{(N-1)^2} \end{bmatrix}, \quad (2.45)$$

$$\Phi_{\mathbf{C}} = \frac{1}{\sqrt{M}} \begin{bmatrix} W_M^0 & W_M^0 & W_M^0 & \cdots & W_M^0 \\ W_M^0 & W_M^1 & W_M^2 & \cdots & W_M^{M-1} \\ W_M^0 & W_M^2 & W_M^4 & \cdots & W_M^{2(M-1)} \\ \cdots & \cdots & \cdots & \ddots & \cdots \\ W_M^0 & W_M^{M-1} & W_M^{2(M-1)} & \cdots & W_M^{(M-1)^2} \end{bmatrix}. \quad (2.46)$$

Zpětnou Fourierovu transformaci lze pak zapsat maticovým vztahem

$$\mathbf{f} = \Phi_{\mathbf{C}}^* \mathbf{F} \Phi_{\mathbf{R}}^*, \quad (2.47)$$

kde $*$ znamená komplexně sdružené matice.

2.3.4 Vlastnosti diskretní Fourierovy transformace

V tomto odstavci uvedeme několik významnějších vlastností diskretní Fourierovy transformace. Jistě nepřekvapí, že většina z nich koresponduje s vlastnostmi Fourierovy transformace spojité. Vlastnosti proto uvedeme pouze ve stručném přehledu. Předpokládejme, že $f(m, n)$, $g(m, n)$ jsou diskretní signály; $F(k, l)$, $G(k, l)$ jsou Fourierovy obrazy těchto signálů; a, b jsou čísla (obecně komplexní). Vybrané vlastnosti diskretní Fourierovy transformace pak můžeme formulovat takto:

a) *Linearita:*

$$\begin{aligned}\mathcal{F}\{af(m, n) + bg(m, n)\} &= a\mathcal{F}\{f(m, n)\} + b\mathcal{F}\{g(m, n)\} \\ &= aF(k, l) + bG(k, l).\end{aligned}\quad (2.48)$$

b) *Periodicita:* Ve výrazu (2.42) pro Fourierovu transformaci jsme předpokládali, že k je z rozsahu $0, 1, 2, \dots, M - 1$ a l z rozsahu $0, 1, 2, \dots, N - 1$. Pokud bychom se pokoušeli vypočítat hodnoty Fourierova obrazu pro k, l mimo uvedené meze, pak analýzou vztahu (2.42) zjistíme, že platí

$$\begin{aligned}F(k, -l) &= F(k, N - l), \\ F(-k, l) &= F(M - k, l), \\ F(-k, -l) &= F(M - k, N - l),\end{aligned}\quad (2.49)$$

$$F(\alpha M + k, \beta N + l) = F(k, l), \quad \alpha, \beta = 0, \pm 1, \pm 2, \dots \quad (2.50)$$

Podobně také z předpisu (2.42) pro zpětnou transformaci vychází

$$\begin{aligned}f(-m, n) &= f(M - m, n), \\ f(m, -n) &= f(m, N - n), \\ f(-m, -n) &= f(M - m, N - n),\end{aligned}\quad (2.51)$$

$$f(\alpha M + m, \beta N + n) = f(m, n), \quad \alpha, \beta = 0, \pm 1, \pm 2, \dots \quad (2.52)$$

Uvedené vlastnosti vyplývají z periodicity exponenciálního členu ve výrazech (2.42), (2.42).

c) *Kruhový posuv:* Necht $f(m, n)$ je periodická funkce s periodou M, N . Pro Fourierovu transformaci funkce $f(m - m_0, n - n_0)$ platí

$$\mathcal{F}\{f(m - m_0, n - n_0)\} = F(k, l) \exp \left[-j2\pi \left(\frac{m_0 k}{M} + \frac{n_0 l}{N} \right) \right]. \quad (2.53)$$

Podobnou vlastnost lze nalézt také ve frekvenční doméně. Je vyjádřena vztahem

$$\mathcal{F}^{-1}\{F(k - k_0, l - l_0)\} = f(m, n) \exp \left[j2\pi \left(\frac{mk_0}{M} + \frac{nl_0}{N} \right) \right]. \quad (2.54)$$

Poznamenejme, že z předpokladu periodicity funkce $f(m, n)$ plyne, že hodnota $f(m - m_0, n - n_0)$ je definována pro každé $m_0 = 0, 1, \dots, M - 1, n_0 = 0, 1, \dots, N - 1$.

d) *Komplexně sdružená hodnota:* Je-li funkce $f(m, n)$ reálná, pak platí

$$F(\alpha M - k, \beta N - l) = F^*(k, l), \quad \alpha, \beta = 1, 2, \dots \quad (2.55)$$

Vlastnost snadno dokážeme, jestliže levou i pravou stranu rovnice (2.55) rozepíšeme podle vztahu (2.42). Pro praktický výpočet diskretní Fourierovy transformace reálných signálů má vlastnost význam. S jejím využitím postačí hodnoty $F(k, l)$ počítat přibližně v polovině bodů.

e) *Konvoluční teorém*: Podobně jako pro spojitou Fourierovu transformaci, platí i pro transformaci diskretní, že Fourierova transformace konvoluce funkcí v prostorové doméně je součinem Fourierových obrazů funkcí v doméně frekvenční a naopak součinu v prostorové doméně odpovídá konvoluce v doméně frekvenční. Konvoluce se v tomto případě uvažuje cyklická. Je tedy

$$\mathcal{F}\{f(m, n) * g(m, n)\} = \frac{1}{\sqrt{MN}} F(k, l) G(k, l), \quad (2.56)$$

$$\mathcal{F}\{f(m, n)g(m, n)\} = \frac{1}{\sqrt{MN}} F(k, l) * G(k, l). \quad (2.57)$$

f) *Korelační teorém*: Definujme nejprve křížovou korelaci R_{fg} a autokorelaci R_{ff} deterministických diskretních funkcí $f(m, n)$, $g(m, n)$ pomocí vztahů

$$R_{fg}(m, n) = \frac{1}{MN} \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} f^*(r-m, s-n)g(r, s), \quad (2.58)$$

$$\begin{aligned} R_{ff}(m, n) &= \frac{1}{MN} \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} f^*(r-m, s-n)f(r, s) \\ &= \frac{1}{MN} \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} f(r+m, s+n)f^*(r, s). \end{aligned} \quad (2.59)$$

Korelační teorém tvrdí následující

$$\mathcal{F}\{R_{fg}(m, n)\} = \frac{1}{\sqrt{MN}} F^*(k, l)G(k, l), \quad (2.60)$$

$$\mathcal{F}\{R_{ff}(m, n)\} = \frac{1}{\sqrt{MN}} F^*(k, l)F(k, l) = \frac{1}{\sqrt{MN}} |F(k, l)|^2. \quad (2.61)$$

Poznamenejme, že podobně jako u konvoluce, tak se i korelace a autokorelace uvažují cyklické (nebo se předpokládá, že funkce $f(m, n)$, $g(m, n)$ jsou periodické).

g) *Parsevalův teorém*: Předpokládejme, že obecně jsou signály $f(x, y)$, $g(x, y)$ komplexní a použijme symbolu * pro označení komplexně sdružené funkce. Parsevalův teorém tvrdí, že platí

$$\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n)g^*(m, n) = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} F(k, l)G^*(k, l). \quad (2.62)$$

Ve speciálním případě, kdy $f(x, y) = g(x, y)$, přechází uvedený výraz na tvar

$$\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |f(m, n)|^2 = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} |F(k, l)|^2. \quad (2.63)$$

2.4 Kosinová transformace

V kosinové transformaci se využívá skutečnosti, že Fourierova transformace funkce, která je symetrická podle obou souřadných os, má pouze reálnou složku. Imaginární složka vyjde všude nulová. Přítomnost pouze reálné složky je výhodná z hlediska paměťové i časové složitosti výpočtů. Symetrii funkce lze zajistit postupem, který dále popíšeme. Mějme diskrétní funkci $f(m, n)$ definovanou nad oblastí $N \times N$ bodů, $m, n \in \{0, 1, \dots, N-1\}$. Tuto funkci doplníme tak, abychom vytvořili funkci symetrickou. To lze provést dvěma způsoby. První z nich vede na funkci definovanou nad oblastí $2N \times 2N$ bodů, $m, n \in \{-N, -N+1, \dots, N-1\}$, a nazývá se proto sudou kosinovou transformací. Druhý vede na funkci definovanou nad oblastí $(2N-1) \times (2N-1)$ bodů, $m, n \in \{-N+1, -N+2, \dots, N-1\}$, a nazývá se proto kosinovou transformací lichou. Podrobněji ukážeme nejprve druhý způsob - tj. lichou kosinovou transformaci. Zavedme symetrickou funkci $\tilde{f}(m, n)$ podle předpisu

$$\tilde{f}(m, n) = \begin{cases} f(m, n), & m \geq 0, n \geq 0 \\ f(-m, n), & m < 0, n \geq 0 \\ f(m, -n), & m \geq 0, n < 0 \\ f(-m, -n), & m < 0, n < 0 \end{cases}. \quad (2.64)$$

Uplatníme-li na funkci $\tilde{f}(m, n)$ diskrétní Fourierovu transformaci, máme

$$F(k, l) = \frac{1}{2N-1} \sum_{m=-N+1}^{N-1} \sum_{n=-N+1}^{N-1} \tilde{f}(m, n) \exp\left(-j2\pi\left(\frac{mk}{2N-1} + \frac{nl}{2N-1}\right)\right),$$

$$k, l = -N+1, \dots, N-1. \quad (2.65)$$

Vezmeme-li v úvahu, že funkce $\tilde{f}(m, n)$ je symetrická, pak po úpravě dostaneme

$$F(k, l) = \frac{1}{2N-1} \sum_{m=-N+1}^{N-1} \sum_{n=-N+1}^{N-1} \tilde{f}(m, n) \cos\left(2\pi\frac{mk}{2N-1}\right) \cos\left(2\pi\frac{nl}{2N-1}\right). \quad (2.66)$$

S ohledem na symetrii dle (2.64) lze tento vztah dále přepsat do tvaru

$$F(k, l) = \frac{4}{2N-1} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} c(m)c(n)f(m, n) \cos\left(2\pi\frac{mk}{2N-1}\right) \cos\left(2\pi\frac{nl}{2N-1}\right), \quad (2.67)$$

kde

$$c(p) = \begin{cases} 1/2, & p = 0 \\ 1, & \text{jinak} \end{cases} \quad (2.68)$$

Funkce c zohledňuje skutečnost, že ve čtyřech symetrických kvadrantech funkce $\tilde{f}(m, n)$ vytvořených podle vztahu (2.64) jsou hodnoty na osách symetrie sdíleny dvěma kvadranty a hodnota v počátku je sdílena čtyřmi kvadranty. Protože funkce $\cos(x)$ je sudá, zjišťujeme na základě vztahu (2.67), že platí $F(k, -l) = F(k, l)$, $F(-k, l) = F(k, l)$, $F(-k, -l) = F(k, l)$ a že tedy i frekvenční spektrum lze rozdělit na čtyři symetrické kvadranty. Hodnoty $F(k, l)$ proto postačí počítat a uchovávat pouze pro nezáporné indexy k, l . Symetrie lze využít i při provádění zpětné kosinové transformace. Z předpisu pro zpětnou Fourierovu transformaci obdržíme pro zpětnou transformaci kosinovou vztah

$$f(m, n) = \frac{4}{2N-1} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} c(k)c(l)F(k, l) \cos\left(2\pi\frac{mk}{2N-1}\right) \cos\left(2\pi\frac{nl}{2N-1}\right). \quad (2.69)$$

Vzorce pro sudou kosinovou transformaci lze odvodit analogicky. Symetrie je v tomto případě dosaženo zrcadlením původního obrazu kolem přímk $x = -1/2$, $y = -1/2$. Aby ve Fourierově transformaci vedla symetrie k anulování členů obsahujících funkci sinus, je zapotřebí posunout počátek souřadné soustavy do bodu o souřadnicích $(-1/2, -1/2)$. Původní souřadnice m, n se tak transformují na hodnoty $(m + 1/2)$, $(n + 1/2)$. V dopředné sudé kosinové transformaci neleží na osách symetrie žádné řady bodů obrazu. Také v případě sudé kosinové transformace platí $F(k, -l) = F(k, l)$, $F(-k, l) = F(k, l)$, $F(-k, -l) = F(k, l)$, a proto i zde postačí počítat a uchovávat pouze jeden kvadrant frekvenčního spektra. Hodnoty $F(k, l)$ ležící na osách symetrie jsou sdíleny sousedními kvadranty spektra. Při výpočtu zpětné transformace lze tuto skutečnost vzít opět v úvahu použitím funkce c . Zpravidla se polovina uvedené korekce realizuje už při dopředné transformaci a polovina při transformaci zpětné. Pro dopřednou a zpětnou sudou kosinovou transformaci tak máme

$$F(k, l) = \frac{4}{2N} c(k)c(l) \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f(m, n) \cos\left[\pi\frac{(2m+1)k}{2N}\right] \cos\left[\pi\frac{(2n+1)l}{2N}\right], \quad (2.70)$$

$$f(m, n) = \frac{4}{2N} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} c(k)c(l)F(k, l) \cos\left[\pi\frac{(2m+1)k}{2N}\right] \cos\left[\pi\frac{(2n+1)l}{2N}\right], \quad (2.71)$$

kde

$$c(p) = \begin{cases} 1/\sqrt{2}, & p = 0 \\ 1, & \text{jinak} \end{cases} \quad (2.72)$$

Na závěr tohoto odstavce poznamenejme, že kosinová transformace nachází významné uplatnění při kompresi obrazových signálů. Využívá se v kompresních metodách JPEG a MPEG, kde se používá kosinové transformace sudé.

Kapitola 3

Stochastický přístup k popisu obrazových signálů

V předchozích kapitolách jsme předpokládali, že obrazové signály popisujeme deterministicky jejich velikostí v jednotlivých bodech obrazu a zpracováváme jednotlivě, nezávisle na jiných obrazových signálech. Tato představa není ovšem vždy účelná. Často chceme například navrhnout nebo vyšetřit parametry systému zpracovávajícího nějakou třídu obrazových signálů, aniž bychom nějaký konkrétní jednotlivý signál měli předem na mysli. V takových případech lze použít přístupu stochastického. Při stochastickém přístupu je obraz charakterizován jistým souborem statistických veličin. V této kapitole popíšeme základní nástroje, které se při stochastickém přístupu používají. Ke studiu kapitoly jsou potřebné základní znalosti z počtu pravděpodobnosti.

3.1 Náhodné pole

Při stochastickém přístupu se signál považuje za *náhodný proces* (jednorozměrné signály) nebo za *náhodné pole* (vícerozměrné signály). Zavedení uvedených pojmů je věnována tato podkapitola. Nechť \mathcal{S} je oblast v m -rozměrném euklidovském prostoru. Nechť R označuje bod této oblasti. Poloha bodu R je popsána vektorem \mathbf{r} . Definujme nyní množinu $\{\mathbf{f}_R\}$ náhodných proměnných tak, že pro každý bod oblasti \mathcal{S} zavedeme jednu *náhodnou proměnnou*. Je-li $m = 1$ (jednorozměrný signál), pak se takto vytvořená množina náhodných proměnných nazývá stochastickým (náhodným) procesem. Je-li $m > 1$, pak se používá termínu náhodné pole nebo termínu vícerozměrný stochastický proces. Nechť Ω označuje množinu všech elementárních jevů a ω_i je i -tý elementární jev. Na náhodný proces (náhodné pole) lze nahlížet ze dvou pohledů: 1) Jedná se o množinu $\{\mathbf{f}_R(\omega_i) | R \in \mathcal{S}\}$ náhodných proměnných. 2) Jedná se množinu $\{f_{\omega_i}(\mathbf{r}) | \omega_i \in \Omega\}$ funkcí nad oblastí \mathcal{S} - jedna funkce pro každý elementární jev ω_i . Pro náhodný proces (náhodné pole) zavedeme značení $\mathbf{f}(\mathbf{r}, \omega_i)$ nebo také pouze stručně $\mathbf{f}(\mathbf{r})$. Při zpracování obrazů pracujeme nejčastěji s dvojrozměrnými náhodnými poli $\mathbf{f}(\mathbf{r})$, kde $\mathbf{r} = (x, y)$. Vezmeme-li v úvahu, že pro každou

hodnotu \mathbf{r} je $\mathbf{f}(\mathbf{r})$ náhodnou proměnnou v obvyklém smyslu, pak není obtížné zavést *distribuční funkci* $P_f(z, \mathbf{r})$ a *hustotu pravděpodobnosti* $p_f(z, \mathbf{r})$ náhodného pole. Jak distribuční funkce, tak hustota pravděpodobnosti jsou v tomto případě také funkcí polohy \mathbf{r} . Distribuční funkce $P_f(z, \mathbf{r})$ je definována obvyklým způsobem - tj. jako pravděpodobnost události $\{\mathbf{f}(\mathbf{r}) \leq z\}$. Je tedy

$$P_f(z, \mathbf{r}) = \mathcal{P}\{\mathbf{f}(\mathbf{r}) \leq z\}. \quad (3.1)$$

Hustotu pravděpodobnosti $p_f(z, \mathbf{r})$ získáme jako obvykle derivováním distribuční funkce:

$$p_f(z, \mathbf{r}) = dP_f(z, \mathbf{r})/dz. \quad (3.2)$$

Uvažujme nyní v oblasti \mathcal{S} n bodů $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n$. Podle dříve uvedeného odpovídají těmto bodům náhodné proměnné $\mathbf{f}(\mathbf{r}_1), \mathbf{f}(\mathbf{r}_2), \dots, \mathbf{f}(\mathbf{r}_n)$. Můžeme tedy pro náhodné pole zavést sdruženou distribuční funkci a sdruženou hustotu pravděpodobnosti n -tého řádu:

$$P_f(z_1, z_2, \dots, z_n, \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) = \mathcal{P}\{\mathbf{f}(\mathbf{r}_1) \leq z_1, \mathbf{f}(\mathbf{r}_2) \leq z_2, \dots, \mathbf{f}(\mathbf{r}_n) \leq z_n\}, \quad (3.3)$$

$$p_f(z_1, z_2, \dots, z_n, \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) = \frac{\partial^n P_f((z_1, z_2, \dots, z_n, \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n)}{\partial z_1 \partial z_2 \dots \partial z_n}. \quad (3.4)$$

3.2 Momenty náhodného pole

Předpokládejme na okamžik, že signál reprezentujeme jeho n vzorky získanými v bodech $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n$. Pokud bychom za této situace měli k dispozici sdruženou hustotu pravděpodobnosti n -tého řádu, byl by takto signál vyčerpávajícím způsobem popsán. Bohužel jsou však postupy založené na sdružené pravděpodobnosti vyšších řádů teoreticky i prakticky obtížně zvládnutelné. Nejlépe jsou zatím propracovány postupy založené na sdružené pravděpodobnosti druhého řádu, které vyšetřují, jak jsou svázány vlastnosti obrazu ve dvou jeho různých bodech. Máme-li pro náhodné pole zavedenu sdruženou hustotu pravděpodobnosti druhého řádu, můžeme definovat jeho střední hodnotu, korelaci a kovarianci. *Střední hodnota* $\mu(\mathbf{r})$ náhodného pole je definována vztahem

$$\mu(\mathbf{r}) = E\{\mathbf{f}(\mathbf{r})\} = \int_{-\infty}^{\infty} z p_f(z, \mathbf{r}) dz. \quad (3.5)$$

Autokorelace $R_{ff}(\mathbf{r}_1, \mathbf{r}_2)$ náhodného pole $\mathbf{f}(\mathbf{r})$ je definována jako korelace náhodných proměnných $\mathbf{f}(\mathbf{r}_1), \mathbf{f}(\mathbf{r}_2)$. Je tedy

$$R_{ff}(\mathbf{r}_1, \mathbf{r}_2) = E\{\mathbf{f}(\mathbf{r}_1)\mathbf{f}(\mathbf{r}_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_1 z_2 p_f(z_1, z_2, \mathbf{r}_1, \mathbf{r}_2) dz_1 dz_2. \quad (3.6)$$

Podobně *autokovariance* $C_{\mathbf{f}}(\mathbf{r}_1, \mathbf{r}_2)$ náhodného pole je definována jako kovariance náhodných proměnných $\mathbf{f}(\mathbf{r}_1), \mathbf{f}(\mathbf{r}_2)$. Odtud máme

$$C_{\mathbf{f}}(\mathbf{r}_1, \mathbf{r}_2) = E \{ [\mathbf{f}(\mathbf{r}_1) - \mu_{\mathbf{f}}(\mathbf{r}_1)] [\mathbf{f}(\mathbf{r}_2) - \mu_{\mathbf{f}}(\mathbf{r}_2)] \}. \quad (3.7)$$

Roznásobením výrazů v hranatých závorkách a uplatněním operace střední hodnoty na každý z takto vzniklých členů lze snadno dokázat, že platí

$$C_{\mathbf{f}}(\mathbf{r}_1, \mathbf{r}_2) = R_{\mathbf{f}}(\mathbf{r}_1, \mathbf{r}_2) - \mu_{\mathbf{f}}(\mathbf{r}_1)\mu_{\mathbf{f}}(\mathbf{r}_2). \quad (3.8)$$

Analogicky můžeme zavést *křížovou korelaci* $R_{\mathbf{fg}}(\mathbf{r}_1, \mathbf{r}_2)$ a *křížovou kovarianci* $C_{\mathbf{fg}}(\mathbf{r}_1, \mathbf{r}_2)$ náhodných polí $\mathbf{f}(\mathbf{r})$ a $\mathbf{g}(\mathbf{r})$:

$$R_{\mathbf{fg}}(\mathbf{r}_1, \mathbf{r}_2) = E \{ \mathbf{f}(\mathbf{r}_1) \mathbf{g}(\mathbf{r}_2) \}, \quad (3.9)$$

$$C_{\mathbf{fg}}(\mathbf{r}_1, \mathbf{r}_2) = E \{ [\mathbf{f}(\mathbf{r}_1) - \mu_{\mathbf{f}}(\mathbf{r}_1)] [\mathbf{g}(\mathbf{r}_2) - \mu_{\mathbf{g}}(\mathbf{r}_2)] \}. \quad (3.10)$$

Náhodné pole $\mathbf{f}(\mathbf{r})$ je nekorelované právě tehdy, jestliže pro každé dva různé vektory $\mathbf{r}_1, \mathbf{r}_2$ platí $C_{\mathbf{f}}(\mathbf{r}_1, \mathbf{r}_2) = 0$. Podobně jsou náhodná pole $\mathbf{f}(\mathbf{r})$ a $\mathbf{g}(\mathbf{r})$ vzájemně nekorelovaná právě tehdy, jestliže pro každé dva vektory $\mathbf{r}_1, \mathbf{r}_2$ platí $C_{\mathbf{fg}}(\mathbf{r}_1, \mathbf{r}_2) = 0$.

3.3 Homogenní a ergodické náhodné pole

Kromě toho, že se v praxi omezujeme na sdružené pravděpodobnosti druhého řádu, zavádíme obvykle ještě další zjednodušující předpoklady. Často se předpokládá homogenita a ergodicita náhodného pole. Náhodné pole se nazývá homogenním, jestliže jeho střední hodnota $\mu(\mathbf{r})$ není závislá na \mathbf{r} (pro všechna \mathbf{r} je střední hodnota stejná) a jestliže jeho autokorelace $R_{\mathbf{f}}(\mathbf{r}_1, \mathbf{r}_2)$ závisí pouze na rozdílu $\mathbf{r}_2 - \mathbf{r}_1$. Uvedené podmínky můžeme zapsat pomocí vztahů

$$\mu_{\mathbf{f}}(\mathbf{r}) = \mu_{\mathbf{f}} = \text{konstanta}, \quad (3.11)$$

$$R_{\mathbf{f}}(\mathbf{r}_1, \mathbf{r}_2) = R_{\mathbf{f}}(\mathbf{r}_1 + \mathbf{r}_0, \mathbf{r}_2 + \mathbf{r}_0), \quad (3.12)$$

kde \mathbf{r}_0 je libovolný vektor. Jestliže ve vztahu (3.12) postupně položíme $\mathbf{r}_0 = -\mathbf{r}_2$ a $\mathbf{r}_0 = -\mathbf{r}_1$, získáme

$$R_{\mathbf{f}}(\mathbf{r}_1, \mathbf{r}_2) = R_{\mathbf{f}}(\mathbf{r}_1 - \mathbf{r}_2, 0) = R_{\mathbf{f}}(0, \mathbf{r}_2 - \mathbf{r}_1). \quad (3.13)$$

S ohledem na to, že z definice autokorelace vyplývá $R_{\mathbf{f}}(\mathbf{r}_1, \mathbf{r}_2) = R_{\mathbf{f}}(\mathbf{r}_2, \mathbf{r}_1)$, a zavedeme-li zápis $R_{\mathbf{f}}(\mathbf{r}, 0) \equiv R_{\mathbf{f}}(\mathbf{r})$, pak ze vztahu (3.13) dostáváme $R_{\mathbf{f}}(\mathbf{r}_1 - \mathbf{r}_2) = R_{\mathbf{f}}(\mathbf{r}_2 - \mathbf{r}_1)$. Proto platí

$$R_{\mathbf{f}}(\mathbf{r}) = R_{\mathbf{f}}(-\mathbf{r}). \quad (3.14)$$

Rozepíšeme-li vektor \mathbf{r} jako $\mathbf{r} = (a, b)$, pak pro výpočet autokorelace $R_{\mathbf{f}}$ homogenního pole máme

$$R_{\mathbf{f}}(\mathbf{r}) = R_{\mathbf{f}}(a, b) = E \{ \mathbf{f}(x + a, y + b) \mathbf{f}(x, y) \} = E \{ \mathbf{f}(x, y) \mathbf{f}(x - a, y - b) \}. \quad (3.15)$$

Poznamenejme ještě, že pojem homogenity náhodného pole je totožný s pojmem stacionarita v širším smyslu, se kterým se můžeme v literatuře také setkat.

Pojem *ergodicity* souvisí s možností pohlížet na homogenní náhodné pole jako na množinu $\{f_{\omega_i}(\mathbf{r})\}$ funkcí, jak bylo vysvětleno v podkapitole 3.1. Necht \mathcal{S} je oblast, nad níž náhodné pole (v našem případě obrazový signál) uvažujeme. Míra této oblasti necht je S . Předpokládejme, že oblast \mathcal{S} je dostatečně veliká - ideálně nekonečná. Stanovme střední hodnotu signálu nad oblastí \mathcal{S} . Máme

$$\mathbf{E} = \lim_{S \rightarrow \infty} \frac{1}{S} \iint_{\mathcal{S}} \mathbf{f}(x, y) dx dy. \quad (3.16)$$

Vztahu (3.16) je třeba rozumět tak, že integraci provádíme jednotlivě pro každý elementární jev ω_i . Pro každý elementární jev obdržíme hodnotu $\mathbf{E}(\omega_i)$. \mathbf{E} je zde tedy náhodnou proměnnou (nezaměňovat s operátorem střední hodnoty). Jestliže náhodná proměnná \mathbf{E} nabývá pro všechny elementární jevy stejné hodnoty, pak takové náhodné pole nazveme ergodickým vzhledem ke střední hodnotě. Podobně můžeme pro náhodné pole $\mathbf{f}(x, y)$ určit

$$\mathbf{R}(a, b) = \lim_{S \rightarrow \infty} \frac{1}{S} \iint_{\mathcal{S}} \mathbf{f}(x - a, y - b) \mathbf{f}(x, y) dx dy. \quad (3.17)$$

Stejně jako v předchozím případě, integrujeme i zde pro každý elementární jev jednotlivě a $\mathbf{R}(a, b)$ je proto náhodné pole. Jestliže pro libovolně zvolené a, b nabývá $\mathbf{R}(a, b)$ pro všechny elementární jevy téže hodnoty, pak náhodné pole $\mathbf{f}(x, y)$ nazveme ergodickým vzhledem k autokorelaci.

3.4 Výkonová spektrální hustota

Výkonová spektrální hustota $G_{\mathbb{F}}(u, v)$ homogenního náhodného pole je Fourierovou transformací jeho autokorelace. Je tedy

$$G_{\mathbb{F}}(u, v) = \mathcal{F}\{R_{\mathbb{F}}(a, b)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{\mathbb{F}}(a, b) \exp[-j2\pi(au + bv)] da db. \quad (3.18)$$

Na základě znalosti výkonové spektrální hustoty lze autokorelaci určit inverzní Fourierovou transformací pomocí vztahu

$$R_{\mathbb{F}}(a, b) = \mathcal{F}^{-1}\{G_{\mathbb{F}}(u, v)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_{\mathbb{F}}(u, v) \exp[j2\pi(au + bv)] du dv. \quad (3.19)$$

Jestliže ve vztahu (3.19) položíme $a = b = 0$, můžeme dokázat následující vlastnost

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_{\mathbb{F}}(u, v) du dv = R_{\mathbb{F}}(0, 0) = \mathbf{E}\{[\mathbf{f}(x, y)]^2\} \geq 0. \quad (3.20)$$

3.5 Lineární operace nad náhodným polem

Nechť $\mathbf{f}(x, y)$ je náhodné pole reprezentující obrazový signál. Náhodné pole je charakterizováno střední hodnotou $\mu_f(x, y)$, autokorelací $R_{ff}(a, b)$ a výkonovou spektrální hustotou $G_{ff}(u, v)$. Na náhodné pole $\mathbf{f}(x, y)$ aplikujeme operátor \mathcal{O} a získáme tak náhodné pole $\mathbf{g}(x, y)$. Je tedy

$$\mathbf{g}(x, y) = \mathcal{O}\{\mathbf{f}(x, y)\}. \quad (3.21)$$

Naším cílem nyní bude pro takto získané náhodné pole $\mathbf{g}(x, y)$ stanovit statistické veličiny, které jej popisují. Hledáme tedy předpis pro střední hodnotu, autokorelaci a výkonovou spektrální hustotu náhodného pole $\mathbf{g}(x, y)$, a to ve vztahu k odpovídajícím veličinám vstupního náhodného pole $\mathbf{f}(x, y)$. Pro stanovení uvedených charakteristik náhodného pole $\mathbf{g}(x, y)$ použijeme závěrů z odstavce 1.2.4. V tomto odstavci jsme odvodili, že je-li \mathcal{O} lineární operace invariantní vůči posuvu, pak operaci $\mathcal{O}\{f(x, y)\}$ je možné provést jako konvoluci $f(x, y) * h(x, y)$, kde $h(x, y) = \mathcal{O}\{\delta(x, y)\}$ je odezva operátoru na Diracův impuls (impulzová charakteristika operátoru). Chápejme nyní podle podkapitoly 3.1 náhodné pole $\mathbf{f}(x, y)$ jako množinu $\{f_{\omega_i}(x, y)\}$ funkcí. Pro výstupní náhodné pole $\mathbf{g}(x, y)$ pak máme

$$\mathbf{g}(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{f}(x - a, y - b) h(a, b) da db. \quad (3.22)$$

Výrazu (3.22) je třeba rozumět tak, že pro každý elementární jev ω_i dává funkce $f_{\omega_i}(x, y)$ množiny $\mathbf{f}(x, y)$ vzniknout právě jedné funkci $g_{\omega_i}(x, y)$ množiny $\mathbf{g}(x, y)$. Je tedy $\mathbf{g}(x, y)$ náhodné pole, jak jsme již dříve předeslali. Lze navíc ukázat, že je-li náhodné pole $\mathbf{f}(x, y)$ homogenní, pak je homogenní i pole $\mathbf{g}(x, y)$. Hledané vztahy pro veličiny charakterizující výstupní náhodné pole uvedeme ve formě tvrzení, která dokážeme. Předpokládáme, že vstupní náhodné pole \mathbf{f} je homogenní.

Pro střední hodnotu μ_g náhodného pole $\mathbf{g}(x, y)$ platí

$$\mu_g = \mu_f H(0, 0), \quad (3.23)$$

kde $H(0, 0)$ je hodnota Fourierova obrazu $H(u, v)$ funkce $h(x, y)$ v bodě $u = 0$, $v = 0$.

Důkaz. Tvrzení dokážeme tak, že na obě strany vztahu (3.22) uplatníme operaci

střední hodnoty. Postupně dostaneme

$$\begin{aligned}
 \mu_g &= E \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{f}(x-a, y-b) h(a, b) da db \right\} \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E \{ \mathbf{f}(x-a, y-b) \} h(a, b) da db \\
 &= \mu_f \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(a, b) da db \\
 &= \mu_f \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(a, b) \exp[-j2\pi(0a + 0b)] da db = \mu_f H(0, 0).
 \end{aligned}$$

□

Pro autokorelaci R_{gg} náhodného pole $\mathbf{g}(x, y)$ platí

$$R_{gg}(a, b) = R_{ff}(a, b) * h(-a, -b) * h(a, b). \quad (3.24)$$

Důkaz. Jestliže obě strany rovnice (3.22) násobíme $\mathbf{g}(x - \alpha, y - \beta)$ a uplatníme operaci střední hodnoty, získáme

$$R_{gg}(\alpha, \beta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{fg}(\alpha - a, \beta - b) h(a, b) da db = R_{fg}(\alpha, \beta) * h(\alpha, \beta). \quad (3.25)$$

Podobně násobíme obě strany rovnice (3.22) výrazem $\mathbf{f}(x + \alpha, y + \beta)$. Uplatněním operace střední hodnoty a substituce $a' = -a$, $b' = -b$ pak postupně máme

$$R_{fg}(\alpha, \beta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{ff}(\alpha + a, \beta + b) h(a, b) da db = R_{ff}(\alpha, \beta) * h(-\alpha, -\beta). \quad (3.26)$$

Dosazením výsledku (3.26) do (3.25) již získáme dokazované tvrzení. □

Pro výkonovou spektrální hustotu náhodného pole $\mathbf{g}(x, y)$ platí

$$G_{gg}(u, v) = G_{ff}(u, v) |H(u, v)|^2. \quad (3.27)$$

Důkaz. Dokazované tvrzení vyplývá z tvrzení předchozího. Stačí provést Fourierovu transformaci rovnice (3.24), využít vlastností (2.29) Fourierovy transformace a vzít v úvahu, že $H^*(u, v)H(u, v) = |H(u, v)|^2$. □

3.6 Statistická analýza obrazových transformací

V této podkapitole se budeme zabývat transformacemi diskrétního signálu z pohledu statistické analýzy. Předpokládáme, že vstupní signál je popsán náhodným polem $\mathbf{f}(m, n)$, kde $m = 0, 1, \dots, M-1$, $n = 0, 1, \dots, N-1$. Transformací obdržíme signál, který je popsán náhodným polem $\mathbf{F}(k, l)$, $k = 0, 1, \dots, M-1$, $l = 0, 1, \dots, N-1$. Ve shodě s výrazem (1.34) máme

$$\mathbf{F}(k, l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \mathbf{f}(m, n) \varphi_{k,l}^*(m, n). \quad (3.28)$$

Pro stručnost přepíšeme uvedený vztah do maticového tvaru (maticový zápis transformace signálu jsme zavedli již v podkapitole 2.2, na rozdíl od uvedené podkapitoly jsou zde však \mathbf{f} , \mathbf{F} náhodné vektory). Dostaneme

$$\mathbf{F} = \Phi \mathbf{f}, \quad (3.29)$$

kde \mathbf{f} , \mathbf{F} jsou sloupcové náhodné vektory rozměru MN a Φ je matice rozměru $MN \times MN$. Úlohou bude určit statistické charakteristiky (střední hodnotu, korelaci a kovarianci) výstupního náhodného pole \mathbf{F} na základě odpovídajících charakteristik vstupního náhodného pole \mathbf{f} . Výraz pro střední hodnotu lze získat uplatněním operace střední hodnoty na obě strany rovnice (3.29). Získáme

$$\mu_{\mathbf{F}} = E\{\mathbf{F}\} = \Phi E\{\mathbf{f}\} = \Phi \mu_{\mathbf{f}}, \quad (3.30)$$

kde $\mu_{\mathbf{f}}$, $\mu_{\mathbf{F}}$ jsou vektory středních hodnot rozměru MN . Pro autokorelaci výstupního náhodného pole dostáváme

$$\mathbf{R}_{\mathbf{FF}} = E\{\mathbf{FF}^{*\top}\} = E\{\Phi \mathbf{ff}^{*\top} \Phi^{*\top}\} = \Phi \mathbf{R}_{\mathbf{ff}} \Phi^{*\top}, \quad (3.31)$$

kde $\mathbf{R}_{\mathbf{ff}}$, $\mathbf{R}_{\mathbf{FF}}$ jsou autokorelační matice vstupního a výstupního náhodného pole. Obě matice mají rozměr $MN \times MN$. Podobně pro autokovarianci odvodíme

$$\begin{aligned} \mathbf{C}_{\mathbf{FF}} &= E\left\{[\mathbf{F} - \mu_{\mathbf{F}}][\mathbf{F} - \mu_{\mathbf{F}}]^{*\top}\right\} \\ &= E\left\{[\Phi \mathbf{f} - \Phi \mu_{\mathbf{f}}][\Phi \mathbf{f} - \Phi \mu_{\mathbf{f}}]^{*\top}\right\} = \Phi \mathbf{C}_{\mathbf{ff}} \Phi^{*\top}, \end{aligned} \quad (3.32)$$

kde $\mathbf{C}_{\mathbf{ff}}$, $\mathbf{C}_{\mathbf{FF}}$ jsou autokovarianční matice vstupního a výstupního náhodného pole. Obě matice mají opět rozměr $MN \times MN$. Poznamenejme nakonec, že má-li transformace nějaké speciální vlastnosti (je-li např. separabilní), pak lze za jistých okolností také pro statistické charakteristiky transformovaného signálu odvodit speciální vztahy, které umožňují efektivnější výpočet než obecné vztahy, které jsme zde právě uvedli. Podrobnější výklad v tomto směru by však pravděpodobně přesáhl skromné možnosti tohoto učebního textu.

Příklad 3.1. Uvažujme homogenní diskrétní náhodné pole $\mathbf{f}(m, n)$. Ukážeme, že pro výkonovou spektrální hustotu platí vztah

$$G_{\mathbf{ff}} = \frac{1}{\sqrt{MN}} \mathbb{E} \{ |\mathbf{F}(k, l)|^2 \}.$$

Z podkapitoly 3.4 víme, že výkonová spektrální hustota je Fourierovou transformací autokorelace homogenního náhodného pole. Je tedy

$$G_{\mathbf{ff}} = \mathcal{F} \{ R_{\mathbf{ff}}(m, n) \}.$$

Z podmínky homogenity plyne (přestože jsou hodnoty $\mathbf{f}(m, n)$ reálné, použijeme od této chvíle k jejich reprezentaci komplexních čísel - imaginární složka bude však nulová)

$$\begin{aligned} R_{\mathbf{ff}}(m, n) &= \mathbb{E} \{ \mathbf{f}^*(r - m, s - n) \mathbf{f}(r, s) \} \\ &= \mathbb{E} \left\{ \frac{1}{MN} \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} \mathbf{f}^*(r - m, s - n) \mathbf{f}(r, s) \right\}. \end{aligned}$$

Na základě linearitě Fourierovy transformace máme

$$\mathcal{F} \{ R_{\mathbf{ff}}(m, n) \} = \mathbb{E} \left\{ \mathcal{F} \left[\frac{1}{MN} \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} \mathbf{f}^*(r - m, s - n) \mathbf{f}(r, s) \right] \right\}.$$

Uplatněním korelačního teorému (2.61) vychází

$$G_{\mathbf{ff}}(k, l) = \mathcal{F} \{ R_{\mathbf{ff}}(m, n) \} \frac{1}{\sqrt{MN}} \mathbb{E} \{ \mathbf{F}(k, l) \mathbf{F}^*(k, l) \} = \frac{1}{\sqrt{MN}} \mathbb{E} \{ |\mathbf{F}(k, l)|^2 \}.$$

Jestliže by náhodné pole $\mathbf{f}(m, n)$ bylo navíc ergodické vzhledem k autokorelaci, pak by dále platilo

$$G_{\mathbf{ff}}(k, l) = \frac{1}{\sqrt{MN}} |F(k, l)|^2.$$

Kapitola 4

Digitalizace obrazových signálů

Při digitálním zpracování obrazu je dosti často cílem zpracovávat obrazy, které vznikají jako obrazy spojité. Spojitá reprezentace může být žádána také pro výstupy zpracování. Například tehdy, jestliže má být výsledek pozorován člověkem. Pro samotné digitální zpracování obrazu v počítači není však spojitá reprezentace vhodná. Vhodnější je reprezentace diskrétní, kdy je obraz reprezentován spočetnou (v praxi konečnou) množinou nějakých hodnot. Z uvedeného důvodu provádíme před digitálním zpracováním převod obrazu do diskrétní podoby. Říkáme, že obraz vzorkujeme. Dále se při praktické reprezentaci obrazů v počítači vzorky někdy uchovávají jako celočíselné hodnoty. V takových případech pak skutečné úrovně signálu na tyto celočíselné hodnoty převádíme, což nazýváme kvantováním. Po zpracování v digitalizované podobě může být naopak požadována rekonstrukce obrazu do podoby spojité. Problematice vzorkování, kvantování a rekonstrukce je věnována tato kapitola.

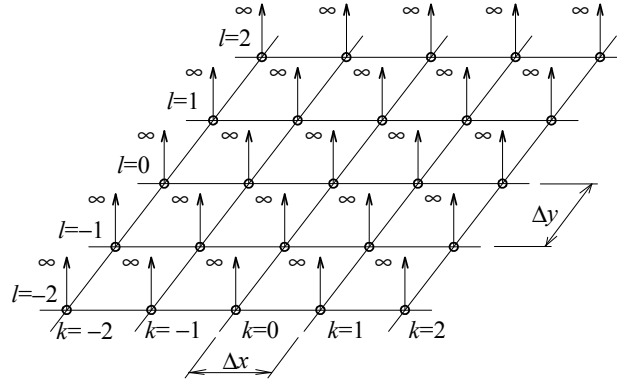
4.1 Vzorkování deterministicky popsaných obrazů

Předpokládejme, že vzorkujeme vstupní obraz deterministicky popsaný spojitou obrazovou funkcí $f_I(x, y)$. Aby vztahy, ke kterým dospějeme, vyšly jednoduché, předpokládáme, že je obraz definován nad nekonečnou oblastí. Výsledek *vzorkování* obrazu je popsán funkcí $f_S(x, y)$. Jako matematický model vzorkování volíme součin vstupního obrazu se vzorkovací funkcí $s(x, y)$. Je tedy

$$f_S(x, y) = f_I(x, y)s(x, y). \quad (4.1)$$

Jako *vzorkovací funkci* nejprve zvolíme nekonečné pole Diracových impulzů navzájem vzdálených o Δx , Δy (obr. 4.1). Tato funkce je popsána předpisem

$$s(x, y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta(x - k\Delta x, y - l\Delta y). \quad (4.2)$$



Obr. 4.1: Vzorkovací funkce ve tvaru nekonečného pole Diracových impulzů

Dosažením vztahu (4.2) do vztahu (4.1) obdržíme pro $f_S(x, y)$ následující předpis

$$f_S(x, y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f_I(x, y) \delta(x - k\Delta x, y - l\Delta y). \quad (4.3)$$

Je zřejmé, že ke stanovení funkce f_S podle předpisu (4.3) stačí znát funkční hodnoty f_I pouze v bodech $(k\Delta x, l\Delta y)$. Při praktické realizaci vzorkování jsou proto právě hodnoty $f_I(k\Delta x, l\Delta y)$ uchovávány. Určeme nyní Fourierovu transformaci $F_S(u, v) = \mathcal{F}\{f_S(x, y)\}$ vzorkovaného obrazu $f_S(x, y)$. Necht $F_I(u, v) = \mathcal{F}\{f_I(x, y)\}$, $S(u, v) = \mathcal{F}\{s(x, y)\}$ jsou Fourierovy transformace funkcí $f_I(x, y)$ a $s(x, y)$. Uplatníme-li na obě strany rovnice (4.1) Fourierovu transformaci a vezmeme-li v úvahu vlastnost (2.31), pak získáme

$$F_S(u, v) = F_I(u, v) * S(u, v). \quad (4.4)$$

Fourierův obraz nekonečného pole Diracových impulzů jsme vypočetli již dříve v příkladě 2.4. Vyšlo

$$S(u, v) = \frac{1}{\Delta x \Delta y} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta\left(u - \frac{k}{\Delta x}, v - \frac{l}{\Delta y}\right). \quad (4.5)$$

Zavedme frekvence u_S, v_S vzorkování jako převrácené hodnoty vzdálenosti vzorků

$$u_S = 1/\Delta x, \quad v_S = 1/\Delta y. \quad (4.6)$$

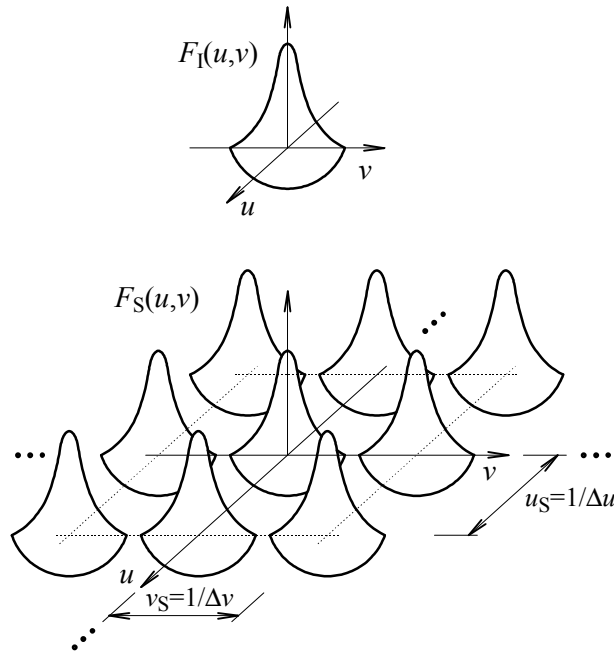
Dosažením vztahů (4.5) a (4.6) do vztahu (4.4) a rozepsáním konvoluce máme

$$F_S(u, v) = u_S v_S \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_I(a, b) \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta(u - a - k u_S, v - b - l v_S) da db. \quad (4.7)$$

Vzájemnou záměnou pořadí integrování a sumace a po úpravě obdržíme

$$F_S(u, v) = u_S v_S \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} F_I(u - ku_S, v - lv_S). \quad (4.8)$$

Nyní jsme došli k významnému závěru: Výraz (4.8) říká, že frekvenční spektrum F_S vzorkovaného obrazu se skládá ze složek, kterými je spektrum F_I obrazu původního. Složky se periodicky opakují ve vzdálenostech u_S, v_S (obr. 4.2) a sčítají. Jak uvidíme později, bude mít pro úspěšnou rekonstrukci obrazu zásadní význam požadavek, aby se jednotlivé složky spektra navzájem nepřekrývaly. Zkoumejme tuto situaci blíže. Předpokládejme, že spektrum vstupního obrazu je omezené - tj., že platí



Obr. 4.2: Spektrum původního a vzorkovaného obrazu

$$F_I(u, v) = 0, \text{ jestliže } |u| > u_I \cup |v| > v_I, \quad (4.9)$$

kde hodnoty u_I, v_I jsou mezní frekvence vstupního obrazu. Z výrazu (4.8) vyplývá (viz. též obr. 4.2), že k překrývání složek spektra nedojde za podmínky, že platí nerovnosti

$$u_S \geq 2u_I, \quad v_S \geq 2v_I. \quad (4.10)$$

Nyní se budeme zabývat otázkou, za jakých podmínek lze ze *vzorkovaného obrazu* $f_S(x, y)$ vytvořit *rekonstruovaný obraz* $f_R(x, y)$ tak, aby byl shodný se vstupním obrazem $f_I(x, y)$. Předpokládáme, že rekonstrukci obrazu bude možné provést pomocí

nějakého filtru, který proto budeme nazývat *filtrem rekonstrukčním*. Očekáváme, že rekonstrukční filtr bude lineární a invariantní vůči posunu. Filtr charakterizujeme odezvou $r(x, y)$ na Diracův impulz (impulzová charakteristika filtru) a Fourierovou transformací $R(u, v)$ této odezvy (frekvenční charakteristika). Tyto funkce však zatím neznáme a chceme je určit.

Z odstavce 1.2.4 víme, že obraz na výstupu rekonstrukčního filtru lze stanovit pomocí konvoluce vzorkovaného obrazu a odezvy filtru na Diracův impulz. Je tedy

$$f_R(x, y) = f_S(x, y) * r(x, y). \quad (4.11)$$

Ve frekvenční doméně pak s ohledem na vlastnost (2.30) Fourierovy transformace máme

$$F_R(u, v) = F_S(u, v)R(u, v). \quad (4.12)$$

Dosadíme-li do vztahu (4.12) ze vztahu (4.8), dostaneme

$$F_R(u, v) = u_S v_S R(u, v) \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} F_I(u - ku_S, v - lv_S). \quad (4.13)$$

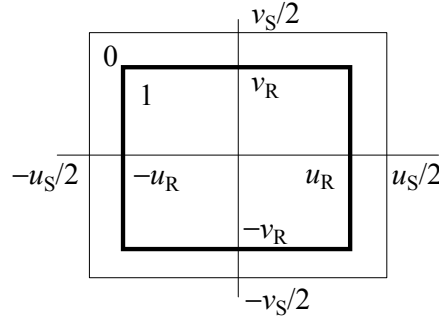
Předpokládejme, že nedojde k překrývání jednotlivých složek spektra, jak jsme již dříve diskutovali (vztahy (4.10)). Ze vztahu (4.13) pak vyplývá, že jestliže navrhujeme rekonstrukční filtr tak, aby odfiltroval ty složky spektra, pro které $k \neq 0$, $l \neq 0$, a složku $k = 0$, $l = 0$ naopak propustil, pak $F_R(u, v) = F_I(u, v)$ a tedy i $f_R(x, y) = f_I(x, y)$, jak jsme pro rekonstrukci požadovali. Docházíme tedy k závěru, že k optimální rekonstrukci by frekvenční charakteristika rekonstrukčního filtru mohla být např. (obr. 4.3)

$$R(u, v) = \begin{cases} 1, & (|u| \leq u_R) \cap (|v| \leq v_R) \\ 0, & \text{jinak} \end{cases}, \quad (4.14)$$

kde $u_I \leq u_R \leq u_S/2$, $v_I \leq v_R \leq v_S/2$. (Přesně vzato dává filtr (4.14) hodnotu $F_R(u, v) = u_S v_S F_I(u, v)$, což však nezmenšuje kvalitu rekonstrukce, protože $u_S v_S$ je známá konstanta.) Zdůrazněme ještě jednou skutečnost, že podmínkou úspěšné rekonstrukce obrazu je splnění požadavku, aby se jednotlivé složky spektra $F_S(u, v)$ vzorkovaného obrazu navzájem nepřekrývaly. Jestliže mezní frekvence u_I , v_I vstupního obrazu považujeme za neměnné, pak lze splnění této podmínky zajistit vhodnou volbou frekvence vzorkování u_S , v_S tak, aby byly splněny nerovnosti (4.10). Uvedené nerovnosti jsou dvojrozměrnou variantou známého Nyquistova vzorkovacího teorému, který říká, že k tomu, aby mohl být signál úspěšně rekonstruován, je nutné, aby vzorkovací frekvence byla alespoň dvojnásobkem nejvyšší frekvence, která je v signálu obsažena.

Při provádění rekonstrukce obrazu může být výhodné vyhnout se Fourierově transformaci, která může být časově složitá. Vztah (4.11) napovídá, jak to udělat. Dosadíme-li do vztahu (4.11) vztah (4.3), pak po provedení konvoluce a úpravě obdržíme předpis

$$f_R(x, y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f_I(k\Delta x, l\Delta y) r(x - k\Delta x, y - l\Delta y). \quad (4.15)$$



Obr. 4.3: Frekvenční charakteristika obdélníkového rekonstrukčního filtru

Abychom mohli vztahu (4.15) prakticky využít, zbývá ještě určit impulzovou charakteristiku $r(x, y)$ rekonstrukčního filtru (4.14). Zpětnou Fourierovou transformací získáme ze vztahu (4.14) předpis

$$r(x, y) = 4u_R v_R \operatorname{sinc}(2xu_R) \operatorname{sinc}(2yv_R). \quad (4.16)$$

Vidíme tedy, že v prostorové doméně lze rekonstrukci obrazu provést podle předpisu (4.15) jako konvoluci vzorků $f_I(k\Delta x, l\Delta y)$ vstupního obrazu s funkcí sinc.

4.2 Zpřesněné modely vzorkování

Zjednodušující předpoklady, které jsme zavedli v podkapitole 4.1, směřovaly k tomu, abychom získali jednoduché a názorné výsledky. Přestože jsou takto získané výsledky užitečné, považujeme za nutné alespoň naznačit možné modifikace uvedeného postupu, jejichž cílem je věrnější popis skutečnosti.

Konečný obraz a konečný počet vzorků: V předchozí podkapitole jsme předpokládali, že obraz, který vzorkujeme, je definován nad nekonečnou oblastí a že vzorkováním získáme nekonečné množství vzorků. V praxi však pracujeme s obrazy, které jsou konečné a konečný musí být také počet vzorků. Zlepšení souhlasu teoretického modelu vzorkování se skutečností by v tomto ohledu bylo možné dosáhnout zavedením vzorkovací funkce ve tvaru konečného pole Diracových impulzů

$$s(x, y) = \sum_{k=-K}^K \sum_{l=-L}^L \delta(x - k\Delta x, y - l\Delta y), \quad (4.17)$$

kde K, L jsou meze pole. Tento postup však bohužel komplikuje další analýzu.

Obecný tvar vzorkovacího pulzu: Až doposud jsme předpokládali, že matematickým modelem získání vzorků je součin obrazu s vzorkovací funkcí. Jako vzorkovací funkci jsme používali funkci složenou z Diracových impulzů. Obecně však mohou mít vzorkovací pulzy i jiný tvar. Jestliže je tvar pulzu popsán funkcí $p(x, y)$, pak můžeme vzorkovací funkci popsat vztahem

$$s(x, y) = \sum_{k=-K}^K \sum_{l=-L}^L p(x - k\Delta x, y - l\Delta y). \quad (4.18)$$

Vzorkování integrací: Senzory, které lze prakticky realizovat, nedokáží sejmout hodnotu v nekonečně malém bodě. Místo toho je sejmuta jakási průměrná hodnota podél citlivé plošky senzoru. Tuto skutečnost lze zohlednit tak, že jako matematický model získávání vzorků volíme integraci

$$f_S(k\Delta x, l\Delta y) = \int_{k\Delta x - \alpha}^{k\Delta x + \alpha} \int_{l\Delta y - \beta}^{l\Delta y + \beta} f_1(x, y) p(x - k\Delta x, y - l\Delta y) dx dy, \quad (4.19)$$

kde α, β jsou poloviny délek stran snímací plošky senzoru, kterou si zde představujeme jako obdélník. Rozměry snímací plošky při tom předpokládáme menší, než je teoretická vzdálenost vzorkovacích bodů. Funkce $p(x, y)$ popisuje průběh citlivosti plošky. Vzorkování integrací má značný význam. To proto, že ať chceme či nechceme, vždy k němu v praxi do určité míry dochází. Ukažme proto další možnou interpretaci vztahu (4.19). Záměnou proměnných převedeme výraz (4.19) na tvar

$$f_S(k\Delta x, l\Delta y) = \int_{-\alpha}^{\alpha} \int_{-\beta}^{\beta} f_1(k\Delta x - a, l\Delta y - b) p(-a, -b) da db. \quad (4.20)$$

Vztah (4.20) je výpočtem konvoluce funkcí $f_1(x, y)$, $p(-x, -y)$. Výsledek konvoluce použijeme jen v bodech $k\Delta x, l\Delta y$ (mimo tyto body předpokládáme funkční hodnotu f_S rovnu 0). Můžeme proto psát

$$f_S(x, y) = [f_1(x, y) * p(-x, -y)] \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta(x - k\Delta x, y - l\Delta y). \quad (4.21)$$

Výraz (4.21) ukazuje, že vzorkování integrací má stejný účinek, jako kdybychom provedli konvoluci vstupního obrazu $f_1(x, y)$ se vzorkovací funkcí $p(-x, -y)$ a pak následně provedli vzorkování polem Diracových impulzů. Jak vyplývá z vlastnosti 2.30 Fourierovy transformace, realizuje konvoluce obrazu se vzorkovací funkcí filtraci obrazu. Ta může být někdy nežádoucí, protože způsobuje degradaci obrazu. Jindy

může být naopak vítaná - např. tehdy, jestliže vzorkování obrazu probíhá na nižší frekvenci, než odpovídá Nyquistovu teorému, a je proto nutné potlačit v obraze vyšší frekvenční složky.

4.3 Aliasing

Ve vztahu (4.8) jsme ukázali, že frekvenční spektrum obrazu po vzorkování se skládá ze složek, které jsou tvořeny frekvenčním spektrem vstupního obrazu. Tyto složky se periodicky opakují ve vzdálenostech u_S, v_S a sčítají. Zjistili jsme také, že k tomu, aby mohl být obraz úspěšně rekonstruován, je nutné, aby se jednotlivé složky spektra navzájem nepřekrývaly. Dosud jsme však neřešili otázku, co se stane v případě, že tato podmínka není splněna. Věnujme se nyní tomuto problému podrobněji. Na základě vztahu (4.8) je možné spektrum obrazu po vzorkování zapsat ve tvaru

$$F_S(u, v) = u_S v_S [F_I(u, v) + F_Q(u, v)] , \quad (4.22)$$

kde $F_I(u, v)$ je spektrum původního obrazu a $F_Q(u, v)$ zahrnuje všechny ty složky spektra $F_S(u, v)$ obrazu po vzorkování, pro které je ve výrazu (4.8) $k \neq 0, l \neq 0$ (složka $k = 0, l = 0$ je zastoupena členem $F_I(u, v)$). Předpokládejme nyní, že se jednotlivé složky spektra navzájem překrývají (obr. 4.4) a že na takto vzorkovaný obraz aplikujeme rekonstrukční filtr s frekvenční charakteristikou

$$R(u, v) = \begin{cases} 1, & (|u| \leq u_S/2) \cap (|v| \leq v_S/2) \\ 0, & \text{jinak} \end{cases} . \quad (4.23)$$

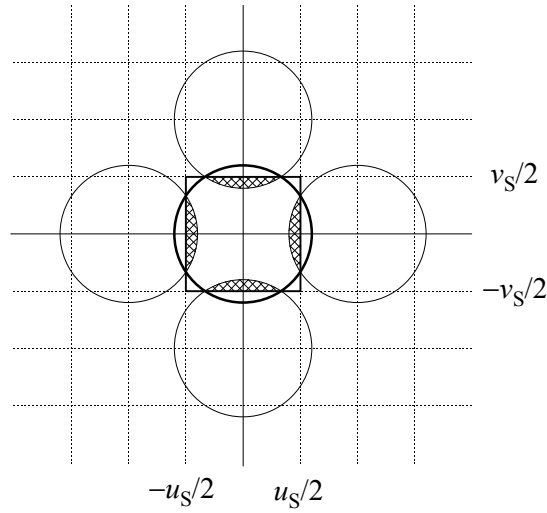
Zvolený rekonstrukční filtr propouští kromě složky $k = 0, l = 0$ (užitečná složka) frekvenčního spektra také určité části složek $k \neq 0, l \neq 0$ (nežádoucí složky). Na obr. 4.4 jsou přenášené části nežádoucích složek vyšrafovány. Naopak si však také můžeme povšimnout, že nejsou přenášeny některé části užitečné složky. Chyba způsobená neúplným přenesením užitečné složky způsobí rozmazání obrazu (potlačení vyšších frekvencí). Naopak chyba způsobená přenosem částí nežádoucích složek dává v obraze vzniknout novým útvarům. Tento jev nazýváme aliasing. Necht $f_{ID}(x, y)$ je obraz degradovaný neúplným přenosem užitečné složky. Vznik aliasingu můžeme matematicky modelovat tak, že na obraz $f_{ID}(x, y)$ superponujeme chybový obraz $a(x, y)$ vzniklý částečným přenosem nežádoucích složek. Je tedy

$$f_R(x, y) = f_{ID}(x, y) + a(x, y) . \quad (4.24)$$

Chybový obraz $a(x, y)$ můžeme popsat vztahem

$$a(x, y) = \mathcal{F}^{-1} \{R(u, v)F_Q(u, v)\} . \quad (4.25)$$

Pokusme se nyní kvantifikovat míru degradace rekonstruovaného obrazu, kterou aliasing způsobí. Ke kvantifikaci použijeme stochastického přístupu. Necht $\mathbf{f}_I(x, y)$ je vstupní náhodné pole a $G_{\mathbf{f}_I}(x, y)$ necht je jeho výkonová spektrální hustota. Na



Obr. 4.4: Vznik aliasingu překrýváním složek spektra vzorkovaného obrazu

míru degradace rekonstruovaného obrazu budeme usuzovat na základě energetické bilance. Energie E_I vstupního obrazového pole je

$$E_I = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_{\text{ff}}(u, v) du dv. \quad (4.26)$$

Z této energie je rekonstrukčním filtrem na výstup přenášeno množství E_{IR} , kde

$$E_{\text{IR}} = \int_{-u_S/2}^{u_S/2} \int_{-v_S/2}^{v_S/2} G_{\text{ff}}(u, v) du dv. \quad (4.27)$$

Rekonstrukčním filtrem je na výstup ovšem také částečně přenášena energie nežádoucích složek. K určení velikosti této energie bychom měli vyšetřovat, jak do oblasti $\langle -u_S/2, u_S/2 \rangle \times \langle -v_S/2, v_S/2 \rangle$ každá z těchto složek přispívá. Není těžké zdůvodnit, že energie nežádoucích složek přenášena v oblasti $\langle -u_S/2, u_S/2 \rangle \times \langle -v_S/2, v_S/2 \rangle$ je stejná jako energie, kterou přenáší užitečná složka mimo tuto oblast (prosíme čtenáře, aby platnost tohoto tvrzení promyslel). Energie E_A nežádoucích složek přenášena rekonstrukčním filtrem na výstup tedy je

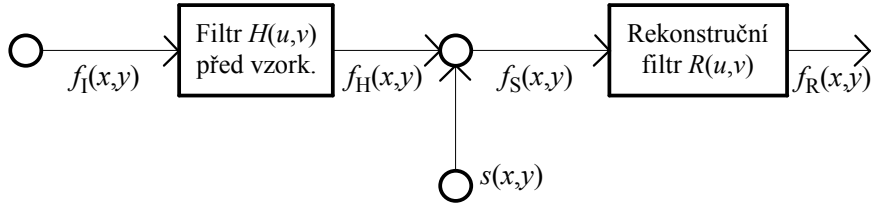
$$E_A = E_I - E_{\text{IR}}. \quad (4.28)$$

Relativní chybu ε_A způsobenou aliasingem lze pak vyjádřit podílem

$$\varepsilon_A = E_A/E_I. \quad (4.29)$$

K omezení jevů způsobených aliasingem můžeme použít filtrace, kterou provedeme ještě před vzorkováním. Cílem filtrace je potlačení vysokých frekvencí v obraze,

čímž budou lépe zajištěny požadavky Nyquistova teorému. Obr. 4.5 znázorňuje matematický model takového postupu: Před vzorkování je zařazen filtr s frekvenční charakteristikou $H(u, v)$. Rekonstrukce je provedena rekonstrukčním filtrem dle vztahu (4.23). Zařazením filtru $H(u, v)$ do řetězce zpracování hrozí ovšem další degradace obrazu. Je tedy nutné zvolit takový filtr, kde omezení jevů aliasingu převažuje nad touto dodatečnou degradací.



Obr. 4.5: Matematický model redukce aliasingu

Nechť $\mathbf{f}_H(x, y)$ je náhodné pole vzniklé aplikací filtru $H(u, v)$ na vstupní náhodné obrazové pole $\mathbf{f}_I(x, y)$ a pole $\mathbf{f}_S(x, y)$ nechť je pole vzniklé vzorkováním pole $\mathbf{f}_H(x, y)$. Označme E_H celkovou energii pole \mathbf{f}_H , a dále označme E_{HR} energii té části užitečné složky pole \mathbf{f}_S , která je propuštěna filtrem $R(u, v)$ na výstup. S využitím vztahu (3.27) máme

$$E_H = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_{\mathbf{f}_I \mathbf{f}_I}(u, v) |H(u, v)|^2 du dv, \quad (4.30)$$

$$E_{HR} = \int_{-u_S/2}^{u_S/2} \int_{-v_S/2}^{v_S/2} G_{\mathbf{f}_I \mathbf{f}_I}(u, v) |H(u, v)|^2 du dv. \quad (4.31)$$

Chybu ε_A způsobenou aliasingem a chybu ε_H způsobenou zařazením filtru $H(u, v)$ lze vyjádřit vztahy

$$\varepsilon_A = (E_H - E_{HR})/E_H, \quad \varepsilon_H = (E_{IR} - E_{HR})/E_{IR}. \quad (4.32)$$

4.4 Praktické metody rekonstrukce obrazu

Úkolem rekonstrukce je získat hodnoty jasu v bodech ležících mimo rastr, v němž jsou k dispozici hodnoty vzorků. Uvedený problém vzniká např. při vykreslování obrazu nebo při provádění geometrických operací s obrazem (kapitola 6). V podkapitole 4.1 jsme ukázali, že rekonstrukci obrazu po vzorkování lze v prostorové doméně

provést konvolucí vzorků obrazu s funkcí sinc (vztahy (4.15)), (4.16)). Protože Fourierovým obrazem funkce sinc je ideální rekonstrukční filtr dle vztahu (4.14), je tento postup za zvolených předpokladů optimální z hlediska přesnosti rekonstrukce. Funkce sinc je však bohužel pro provádění konvoluce poněkud nepříjemná. Její hodnoty jsou téměř všude nenulové, a proto do každého bodu rekonstruovaného obrazu teoreticky přispívají všechny vzorky. Tato skutečnost vede k vysoké časové složitosti rekonstrukce. Navíc je přesné splnění předpokladů, za nichž byl ideální rekonstrukční filtr odvozen, prakticky sotva možné. V praxi se proto často používají filtry, které sice nejsou ideálními rekonstrukčními filtry, ale zato umožňují rychlý výpočet konvoluce. Rychlosti je dosahováno tím, že je impulzová charakteristika filtru nenulová jen na malé oblasti, takže do každého bodu rekonstruovaného obrazu přispívá pouze omezený počet vzorků. Frekvenční charakteristika takových filtrů se však zpravidla dosti odchyľuje od charakteristiky rekonstrukčního filtru ideálního, což teoreticky vede k nižší kvalitě rekonstrukce. Uvedme několik příkladů impulzních a odpovídajících frekvenčních charakteristik častěji používaných rekonstrukčních filtrů:

1) sinc:

$$r(x, y) = u_S v_S \operatorname{sinc}(u_S x) \operatorname{sinc}(v_S y), \quad (4.33)$$

$$R(u, v) = \begin{cases} 1, & (|u| \leq u_S/2) \cap (|v| \leq v_S/2) \\ 0, & \text{jinak} \end{cases}; \quad (4.34)$$

2) čtverec:

$$r_0(x, y) = \begin{cases} 1, & (-\Delta x/2 < x \leq \Delta x/2) \cap (-\Delta y/2 < y \leq \Delta y/2) \\ 0, & \text{jinak} \end{cases}, \quad (4.35)$$

$$R_0(u, v) = \Delta x \Delta y \operatorname{sinc}(u \Delta x) \operatorname{sinc}(v \Delta y); \quad (4.36)$$

3) trojúhelník:

$$r_1(x, y) = r_0(x, y) * r_0(x, y), \quad R_1(u, v) = R_0^2(u, v); \quad (4.37)$$

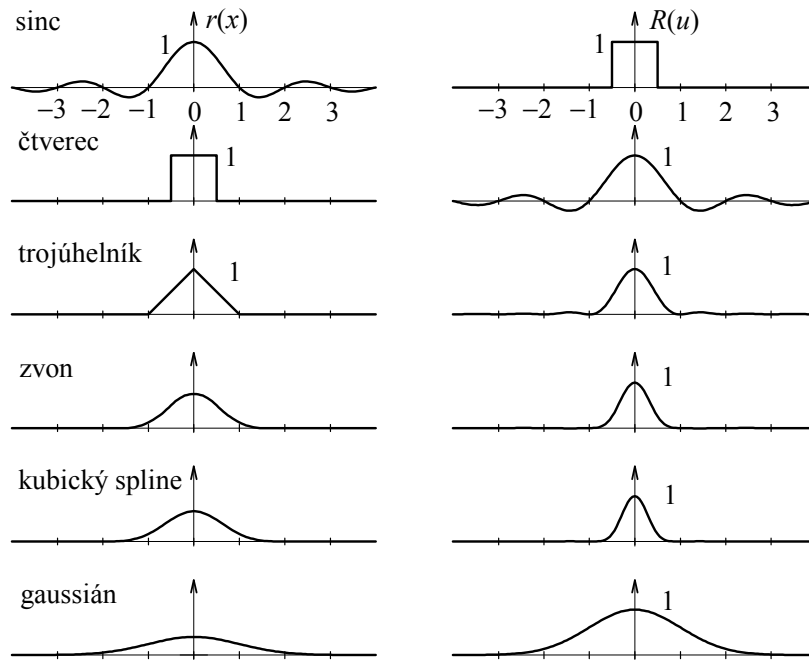
4) zvon:

$$r_2(x, y) = r_0(x, y) * r_1(x, y), \quad R_2(u, v) = R_0^3(u, v); \quad (4.38)$$

5) kubický B-spline:

$$r_3(x, y) = r_0(x, y) * r_2(x, y), \quad R_3(u, v) = R_0^4(u, v); \quad (4.39)$$

6) gaussian:

Obr. 4.6: Průběhy častěji používaných rekonstrukčních filtrů ($\Delta x = u_S = 1$)

$$r(x, y) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{x^2 + y^2}{2\sigma^2}\right], \quad R(u, v) = \exp\left[-\frac{\sigma^2(u^2 + v^2)}{2}\right]. \quad (4.40)$$

Průběhy uvedených funkcí (v jednorozměrné variantě) jsou uvedeny na obr. 4.6. Poznamenejme, že kvantifikaci věrnosti rekonstrukce, které lze zvoleným filtrem dosáhnout, je možné provést podobně jako v odstavci 4.3 energetickou bilancí. K hodnocení může sloužit množství energie, které zvolený rekonstrukční filtr přenáší v pásmu, kde filtr ideální má frekvenční charakteristiku rovnu nule, a také rozdíl v množství energií přenášených zvoleným a ideálním filtrem v pásmu, kde filtr ideální má frekvenční charakteristiku rovnu jedné.

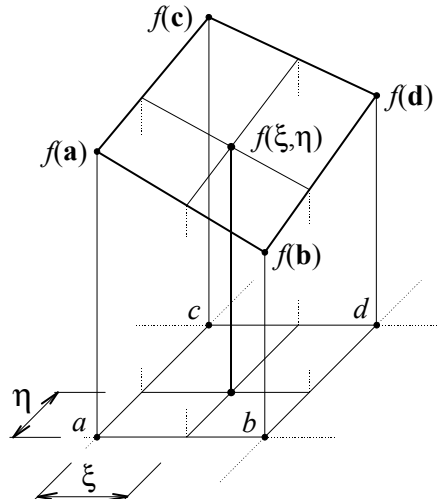
V praxi se rekonstrukce dosti často provádí interpolací. Nejčastěji se používá interpolace bilineární nebo bikubické. Řekněme, že hodnota jasové funkce má být stanovena v bodě o souřadnicích x, y a že $\Delta x = \Delta y = 1$. Zavedme souřadnice ξ, η ($0 \leq \xi, \eta < 1$) pomocí vztahů

$$\xi = x - [x], \quad \eta = y - [y]. \quad (4.41)$$

Dále pro stručnost zápisu položíme

$$\mathbf{a} = ([x], [y]), \quad \mathbf{b} = ([x] + 1, [y]), \quad (4.42)$$

$$\mathbf{c} = (\lfloor x \rfloor, \lfloor y \rfloor + 1), \quad \mathbf{d} = (\lfloor x \rfloor + 1, \lfloor y \rfloor + 1).$$



Obr. 4.7: Bilineární interpolace jasu

Nechť symboly a, b, c, d označují body, které jsou reprezentovány vektory $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$. Bilineární interpolace je popsána následujícím vztahem (obr. 4.7)

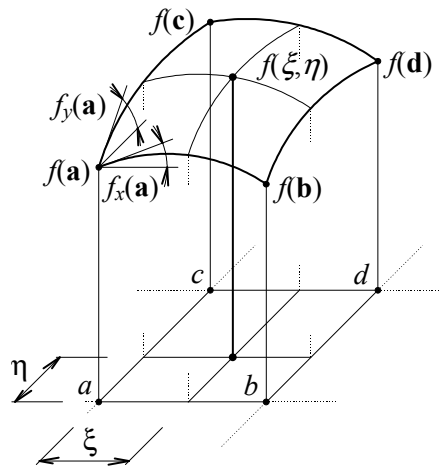
$$\begin{aligned} f(x, y) &= f(\lfloor x \rfloor + \xi, \lfloor y \rfloor + \eta) & (4.43) \\ &= [(1 - \xi), \xi] \begin{bmatrix} f(\mathbf{a}) & f(\mathbf{c}) \\ f(\mathbf{b}) & f(\mathbf{d}) \end{bmatrix} \begin{bmatrix} (1 - \eta) \\ \eta \end{bmatrix} \\ &= (1 - \xi)(1 - \eta)f(\mathbf{a}) + \xi(1 - \eta)f(\mathbf{b}) + (1 - \xi)\eta f(\mathbf{c}) + \xi\eta f(\mathbf{d}). \end{aligned}$$

Poznamenejme, že vztah (4.43) dává pro hodnotu $f(x, y)$ tentýž výsledek jako konvoluce s filtrem $r_1(x, y)$ dle vztahu (4.37). Vztah (4.43) má ovšem velmi názornou interpretaci a také v programech jej lze velmi snadno realizovat.

Při vyšších nárocích na kvalitu interpolace lze použít interpolace bikubické (obr. 4.8). Výpočet lze provést pomocí vztahu

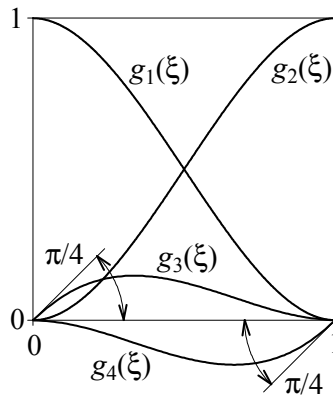
$$\begin{aligned} f(\lfloor x \rfloor + \xi, \lfloor y \rfloor + \eta) &= & (4.44) \\ [g_1(\xi), g_2(\xi), g_3(\xi), g_4(\xi)] & \begin{bmatrix} f(\mathbf{a}) & f(\mathbf{c}) & f_y(\mathbf{a}) & f_y(\mathbf{c}) \\ f(\mathbf{b}) & f(\mathbf{d}) & f_y(\mathbf{b}) & f_y(\mathbf{d}) \\ f_x(\mathbf{a}) & f_x(\mathbf{c}) & 0 & 0 \\ f_x(\mathbf{b}) & f_x(\mathbf{d}) & 0 & 0 \end{bmatrix} \begin{bmatrix} g_1(\eta) \\ g_2(\eta) \\ g_3(\eta) \\ g_4(\eta) \end{bmatrix}. \end{aligned}$$

Funkce g_1, g_2, g_3, g_4 jsou definovány následujícími vztahy (obr. 4.9)



Obr. 4.8: Bikubická interpolace jasu

$$\begin{aligned} g_1(\xi) &= 1 - 3\xi^2 + 2\xi^3, & g_2(\xi) &= 3\xi^2 - 2\xi^3, \\ g_3(\xi) &= \xi - 2\xi^2 + \xi^3, & g_4(\xi) &= -\xi^2 + \xi^3. \end{aligned} \quad (4.45)$$

Obr. 4.9: Funkce $g_1(\xi)$, $g_2(\xi)$, $g_3(\xi)$, $g_4(\xi)$

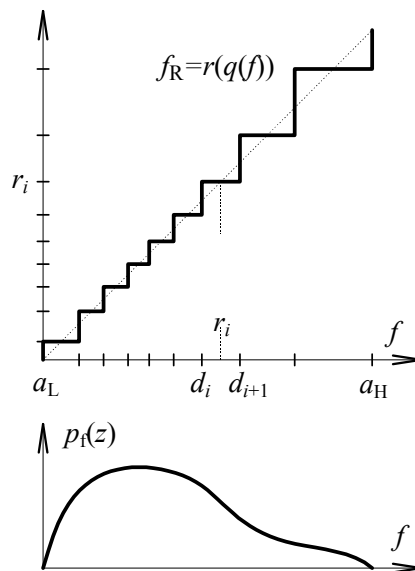
Hodnoty f_x , f_y jsou hodnoty derivací (podle x resp. podle y) jasové funkce v odpovídajícím bodě. Hodnoty derivací nahradíme diferencemi. Např.

$$\begin{aligned} f_x(\mathbf{a}) &= \frac{f(\lfloor x \rfloor + 1, \lfloor y \rfloor) - f(\lfloor x \rfloor - 1, \lfloor y \rfloor)}{2}, \\ f_y(\mathbf{a}) &= \frac{f(\lfloor x \rfloor, \lfloor y \rfloor + 1) - f(\lfloor x \rfloor, \lfloor y \rfloor - 1)}{2}. \end{aligned} \quad (4.46)$$

Na krajích obrazu, kde nejsou hodnoty požadované vztahem (4.46) k dispozici, lze použít nesymetrických diferenčních vzorců.

4.5 Kvantování obrazových signálů

Při praktické reprezentaci signálů v počítači je často potřebné uchovávat jednotlivé vzorky signálu v proměnných s omezeným rozsahem (byte, slovo atd.). V takových případech provádíme převod skutečných úrovní signálu na hodnoty, které lze v proměnné zvoleného typu uchovávat, což nazýváme *kvantováním*. Na konci řetězce digitálního zpracování je naopak často potřebné obraz na základě jeho kvantované reprezentace opět zrekonstruovat tak, aby co nejlépe odpovídal obrazu teoretickému. V tomto odstavci formulujeme uvedenou úlohu poněkud podrobněji a naznačíme také její možné řešení.



Obr. 4.10: Kvantování jasu

Zvolme v obraze nějaký libovolný bod. Necht f je hodnota jasu v tomto bodě. Předpokládejme, že f vždy nabývá reálných hodnot $a_L \leq f < a_H$, kde a_L , a_H značí známou dolní a horní mez jasu. Proměnné, které používáme k uchovávání vzorků, necht umožňují uložení celých čísel v rozsahu 0 až $I - 1$. Přechod od reálné hodnoty f k celočíselné hodnotě $i \in \{0, 1, \dots, I - 1\}$ lze uskutečnit podle předpisu $i = q(f)$, který nazýváme kvantováním. Problémem je nalézt optimální kvantovací funkci q . Kvantovací funkce bývá obvykle zkonstruována následujícím postupem: Zvolme $I - 1$ hodnot d_i , $i = 1, 2, \dots, I - 1$, které interval (a_L, a_H) rozdělí na I podintervalů. Předpokládejme, že platí $d_i < d_{i+1}$ a zavedme dále označení $a_L = d_0$, $a_H = d_I$. Kvantovací funkce q zobrazuje všechny hodnoty jasu z intervalu

$\langle d_i, d_{i+1} \rangle$ na celé číslo i . Problém nalezení kvantovací funkce se v tomto případě redukuje na nalezení $I - 1$ hodnot d_1, d_2, \dots, d_{I-1} . Označme f_R hodnotu jasu po rekonstrukci. K rekonstrukci použijeme funkce $f_R = r(i)$. Rekonstrukční funkce r může být v praxi snadno realizována tabulkou. Funkce r zobrazuje celočíselnou hodnotu i na reálnou hodnotu r_i . V úloze o kvantování hledáme pro daný rozsah I a pro jistou třídu předpokládaných obrazových signálů hodnoty d_i a r_i tak, aby rekonstruované hodnoty jasu v jistém smyslu co nejlépe odpovídaly jasu původnímu (obr. 4.10).

K řešení úlohy, kterou jsme právě formulovali, můžeme použít stochastického přístupu. Považujme vzorek jasu za náhodnou proměnnou s hustotou pravděpodobnosti $p_f(z)$ (obr. 4.10, připomeňme, že hodnota $p_f(z) dz$ udává pravděpodobnost jevu, že jas padne do intervalu $(z, z + dz)$). Označme \mathbf{f} , \mathbf{f}_R náhodné proměnné popisující velikost původního a rekonstruovaného jasu. Hodnoty d_i, r_i nalezneme tak, aby byla minimalizována kvantizační chyba $\varepsilon = E\{(\mathbf{f}_R - \mathbf{f})^2\}$. Uvedený vztah vychází z představy, že pro jasy, které se v obraze vyskytují málo, lze tolerovat větší diferenci $f_R - f$ než pro jasy, které se vyskytují často. Použitím operace střední hodnoty je zajištěno, že se difference $f_R - f$ na celkové chybě podílí úměrně hustotě pravděpodobnosti jasu p_f . Úpravami postupně dostáváme

$$\begin{aligned} \varepsilon &= E\{[\mathbf{f}_R - \mathbf{f}]^2\} = E\{[r(q(\mathbf{f})) - \mathbf{f}]^2\} \\ &= \int_{a_L}^{a_H} [r(q(z)) - z]^2 p_f(z) dz = \sum_{i=0}^{I-1} \int_{d_i}^{d_{i+1}} (r_i - z)^2 p_f(z) dz. \end{aligned} \quad (4.47)$$

Jestliže je počet I úrovní dostatečně velký, můžeme často hustotu pravděpodobnosti $p_f(z)$ uvnitř jednoho kvantovacího intervalu $\langle d_i, d_{i+1} \rangle$ považovat za konstantní, a to $p_f(r_i)$, protože zajisté bude $d_i \leq r_i \leq d_{i+1}$. Ze vztahu (4.47) pak dostáváme

$$\begin{aligned} \varepsilon &= \sum_{i=0}^{I-1} p_f(r_i) \int_{d_i}^{d_{i+1}} (r_i - z)^2 dz \\ &= \frac{1}{3} \sum_{i=0}^{I-1} p_f(r_i) [(d_{i+1} - r_i)^3 - (d_i - r_i)^3]. \end{aligned} \quad (4.48)$$

Z podmínky minima $\frac{\partial \varepsilon}{\partial r_i} = 0$ stanovíme

$$r_i = \frac{d_{i+1} + d_i}{2}. \quad (4.49)$$

Dosazením vztahu (4.49) do (4.48) vyjde

$$\varepsilon = \frac{1}{12} \sum_{i=0}^{I-1} p_f(r_i) (d_{i+1} + d_i)^3. \quad (4.50)$$

Optimální hodnoty úrovní d_i lze nalézt minimalizací hodnoty ε . Řešení této optimalizační úlohy je poněkud nepříjemné. V [14] se jako přibližné řešení uvádí vztah

$$d_i = \frac{(a_H - a_L) \int_{a_L}^{a_i} [p_f(z)]^{-1/3} dz}{\int_{a_L}^{a_H} [p_f(z)]^{-1/3} dz} + a_L, \quad \text{kde } a_i = \frac{i(a_H - a_L)}{I} + a_L. \quad (4.51)$$

Ze vztahu (4.51) je zřejmé, že jestliže je rozdělení pravděpodobnosti rovnoměrné, vychází také rozhodovací úrovně d_i rozmístěny pravidelně.

Komplikace nastávají v případě, kdy je počet kvantovacích úrovní malý, takže hustotu pravděpodobnosti na kvantovacím intervalu nelze považovat za konstantní. V takovém případě není možné provést zjednodušení použité ve výrazu (4.48), ale je nutné pracovat s výrazem (4.47). Za povšimnutí stojí také problém kvantování v barevných obrazech. Jak víme, je barva často reprezentována trojicí R, G, B intenzit základních barevných složek. Vlivem kvantování intenzity barevných složek může u barevných obrazů dojít k nepřesnostem nejen v intenzitě, ale i v barevném odstínu.

Kapitola 5

Filtrace obrazových signálů

Filtrace je jednou ze základních metod úpravy signálů (včetně signálů reprezentujících obrazy) spočívající v úpravě kmitočtového spektra. V principu lze filtraci provádět dvěma postupy: nerekurzivně a rekurzivně. Uvažujme prostor diskretních signálů. Nerekurzivní filtry vytvářejí výstupní signál v každém bodě jako lineární kombinaci vzorků signálu vstupního. Impulzová charakteristika nerekurzivních filtrů obsahuje vždy jen konečný počet nenulových hodnot, protože jinak by je nebylo možné prakticky realizovat. Nerekurzivní filtry jsou proto také označovány jako filtry FIR (finite-extent impulse response). Rekurzivní filtry naproti tomu předepisují vztah mezi vstupním a výstupním signálem obecněji: výstupní signál v každém bodě je lineární kombinací vzorků nejen signálu vstupního, ale také právě určovaného signálu výstupního. Jako rekurzivní lze realizovat také filtry, jejichž impulzová charakteristika obsahuje nekonečný počet nenulových hodnot. Rekurzivní filtry proto také bývají označovány jako filtry IIR (infinite-extent impulse response). Problematice nerekurzivní i rekurzivní filtrace je věnována tato kapitola. Kromě teoretických základů uvedeme také některé praktické metody návrhu filtrů, včetně metod využívajících stochastického přístupu.

5.1 Nerekurzivní filtrace

Nechť $f(m, n)$ je diskretní vstupní signál a $h(m, n)$ nechť je impulzová charakteristika filtru. Nerekurzivní filtrací rozumíme úpravu signálu podle předpisu

$$g(m, n) = f(m, n) * h(m, n). \quad (5.1)$$

Provedeme-li Fourierovu transformaci rovnice (5.1), získáváme vztah

$$G(k, l) = F(k, l) * H(k, l). \quad (5.2)$$

Z rovnice (5.2) je patrné, že filtrace provádí úpravu frekvenčního spektra signálu. Tato představa bývá ostatně dosti často primární. Často požadujeme, aby filtr modifikoval frekvenční spektrum signálu jistým předepsaným způsobem, tj. aby měl jistou požadovanou frekvenční charakteristiku.

5.1.1 Realizace nerekurzivního filtru v prostorové a frekvenční doméně

V praxi se výpočet filtrace provádí jak v doméně prostorové podle vztahu (5.1), tak v doméně frekvenční podle vztahu (5.2). Jedním z faktorů, které ovlivňují volbu postupu, je časová složitost výpočtu. Uvedeme příklad časové rozvahy. Předpokládejme, že máme provést filtraci dle vztahu (5.1) a že funkce f , h jsou nenulové nad čtvercovými oblastmi o rozměrech $M \times M$ a $R \times R$ bodů. Předpokládejme, že vstupní signál f i impulzová charakteristika h filtru jsou reálné, což bývá v praxi často splněno. Časovou složitost budeme měřit počtem reálných násobení. Protože při praktické realizaci konvoluce máme vypočítat M^2 hodnot a protože výpočet každé z nich vyžaduje R^2 násobení, dostáváme pro časovou složitost C_S filtrace realizované v prostorové doméně vztah

$$C_S = M^2 R^2. \quad (5.3)$$

Při provádění filtrace ve frekvenční doméně je třeba provést Fourierovu transformaci vstupního signálu, pak součin na pravé straně vztahu (5.2) a konečně zpětnou Fourierovu transformaci. Předpokládáme, že frekvenční charakteristiku $H(k, l)$ použitého filtru známe předem a že ji není nutné počítat. Protože přechod od vztahu (5.1) ke vztahu (5.2) předpokládá použití cyklické konvoluce, převedeme výpočet konvoluce lineární (v praktických aplikacích se konvoluce ve vztahu (5.1) dosti často uvažuje lineární) na výpočet konvoluce cyklické tak, jak jsme již dříve popsali v odstavci 1.2.5. Zavedme označení $Q = M + R - 1$. Předpokládejme, že časová složitost rychlé Fourierovy transformace je $3Q^2 \log_2 Q/4$ komplexních násobení - tj. $3Q^2 \log_2 Q$ násobení reálných (jedno násobení komplexní vyžaduje čtyři násobení reálná). Celková časová složitost C_F filtrace realizované ve frekvenční doméně měřená počtem reálných násobení tedy je

$$C_F = Q^2(3 \log_2 Q + 4 + 3 \log_2 Q) = Q^2(4 + 6 \log_2 Q). \quad (5.4)$$

Porovnáním vztahů (5.3) a (5.4) zjišťujeme, že provedení filtrace ve frekvenční doméně je výhodnější, jestliže platí

$$R > \sqrt{4 + 6 \log_2 Q}. \quad (5.5)$$

Např. pro $Q = 1024$ vychází, že filtrace ve frekvenční doméně je při uvedených kriteriích výhodnější pro $R > 8$.

5.1.2 Filtry s nulovou fází

Úpravu frekvenčního spektra intuitivně chápeme tak, že se filtrací mění amplituda jednotlivých frekvenčních složek signálu, nikoli však jejich fáze. Předpokládáme tedy, že frekvenční charakteristika filtru je reálná (obecněji můžeme na takový filtr pohlízet jako na filtr, jehož frekvenční charakteristika má nulové imaginární složky,

a proto jej nazýváme filtrem s nulovou fází). Oprávněnost uvedeného požadavku je stvrzována praktickými zkušenostmi - při zpracování obrazu filtry, jejichž frekvenční charakteristika má nenulové imaginární složky, lze subjektivně zaznamenat nežádoucí degradace obrazu. Podmínku nulové fáze filtru ve frekvenční doméně můžeme zapsat pomocí vztahu

$$H(k, l) = H^*(k, l). \quad (5.6)$$

S využitím předpisu (2.42) pro Fourierovou transformaci odtud snadno získáváme podmínku

$$h(m, n) = h^*(-m, -n). \quad (5.7)$$

Vidíme tedy, že podmínka nulové fáze frekvenční charakteristiky filtru vyžaduje, aby impulzová charakteristika filtru byla centrována kolem počátku podle vztahu (5.7). Při filtraci obrazů je někdy možné volit impulzovou charakteristiku filtru na základě intuice. I v takovém případě lze ovšem z výše popsaných důvodů doporučit respektování vztahu (5.7).

5.1.3 Návrh filtru s využitím výřezové funkce

Často chceme v prostorové doméně realizovat filtr, jehož požadovanou impulzovou charakteristiku (označme ji např. $i(m, n)$) sice známe, avšak její praktické použití je nevýhodné, protože je nenulová nad příliš rozsáhlou oblastí, což by vedlo k vysoké časové složitosti výpočtu. V těchto případech lze vytvořit aproximaci $h(m, n)$ ideální impulzové charakteristiky $i(m, n)$ jejím zkrácením pomocí výřezové (okenní) funkce $w(m, n)$ podle předpisu

$$h(m, n) = i(m, n)w(m, n). \quad (5.8)$$

Návrh filtru nyní spočívá ve volbě vhodného typu výřezové funkce w a ve volbě velikosti výřezu tak, aby se vlastnosti filtru dostatečně shodovaly s vlastnostmi požadovaného filtru ideálního. Při volbě výřezové funkce se zpravidla požaduje splnění následujících vlastností:

$$w(m, n) = w^*(-m, -n), \quad (5.9)$$

$$w(m, n) = w_1(m)w_2(n). \quad (5.10)$$

Místo separability dle vztahu (5.10) může být někdy požadována funkce rotačně symetrická, což je zajištěno předpisem

$$w(m, n) = w_0(\sqrt{m^2 + n^2}). \quad (5.11)$$

Splnění uvedených vlastností většinou usnadňuje realizaci následného výpočtu. Splnění vlastnosti (5.9) je předpokladem pro vytvoření filtru s nulovou fází. Separabilita funkce podle (5.10) ulehčuje např. výpočet Fourierovy transformace. Je totiž

$$\mathcal{F}\{w(m, n)\} = \mathcal{F}\{w_1(m)w_2(n)\} = \mathcal{F}\{w_1(m)\} \mathcal{F}\{w_2(n)\} = W_1(k)W_2(l). \quad (5.12)$$

V praxi se používá mnoha typů výřezových funkcí. Zde uvedeme pouze některé z nich (dále uvedené funkce mohou být použity na místě funkcí w_1 , w_2 , w_0 ; Q označuje velikost výřezu):

Obdélník:

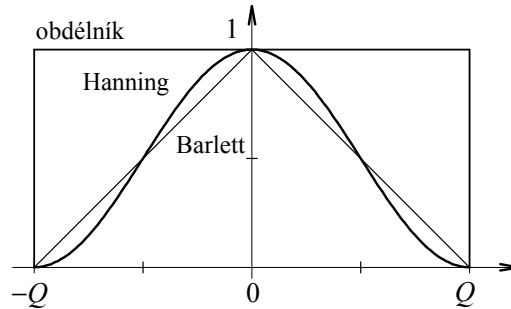
$$w(n) = \begin{cases} 1, & |n| < Q \\ 0, & \text{jinak} \end{cases}, \quad (5.13)$$

Barlettova funkce:

$$w(n) = \begin{cases} \left(1 - \frac{|n|}{Q}\right), & |n| < Q \\ 0, & \text{jinak} \end{cases}, \quad (5.14)$$

Hanningova funkce:

$$w(n) = \begin{cases} \frac{1}{2} \left[1 + \cos\left(\pi \frac{n}{Q}\right)\right], & |n| < Q \\ 0, & \text{jinak} \end{cases}. \quad (5.15)$$



Obr. 5.1: Průběhy některých výřezových funkcí

Průběhy uvedených funkcí jsou znázorněny na obr. 5.1. Průběh frekvenční charakteristiky $\mathcal{F}\{i(m, n)w(m, n)\}$ lze použít pro posouzení kvality filtru. Uplatníme-li na obě strany rovnice (5.8) Fourierovu transformaci, pak pro frekvenční charakteristiku získáme vztah

$$H(k, l) = I(k, l) * W(k, l), \quad (5.16)$$

kde I je frekvenční charakteristika ideálního filtru a W je Fourierův obraz výřezové funkce.

5.1.4 Návrh optimálního filtru

V tomto odstavci se budeme věnovat praktické metodě návrhu impulzové charakteristiky $h(m, n)$ filtru. Uvažujme obrazy rozměru $M \times N$. Teoreticky má tentýž rozměr i impulzová charakteristika filtru. Budeme však požadovat, aby hodnoty $h(m, n)$ byly nenulové jen nad zvolenou a relativně malou oblastí, což je nezbytné z hlediska dosažení přijatelné časové složitosti výpočtu pomocí konvoluce v prostorové doméně

(z podkapitoly 1.2.5 víme, že je pak výhodné nenulové hodnoty impulzové charakteristiky filtru vhodně přeskládat a odpovídajícím způsobem modifikovat praktický předpis pro výpočet konvoluce). Označme Ω množinu dvojic (m, n) , kde mají být hodnoty $h(m, n)$ nalezeny. Předpokládejme, že Ω je obdélníkovou oblastí rozměru $R \times S$. Dále předpokládejme, že známe požadovanou frekvenční charakteristiku filtru. Těto ideální frekvenční charakteristiky patrně nebude možné s ohledem na dříve zavedené omezení dosáhnout přesně. Nechť I je ideální a H skutečná frekvenční charakteristika filtru (obě charakteristiky předpokládáme reálné). Nalezený filtr bude tím lepší, čím menší bude chyba

$$\varepsilon = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} W_{k,l} [H(k, l) - I(k, l)]^2, \quad (5.17)$$

kde $W_{k,l}$ jsou váhové konstanty zohledňující význam dílčí chyby pro jednotlivé frekvenční složky. Tyto konstanty volíme. Vždy však musí být $W_{k,l} > 0$. Za optimální považujeme takový filtr, který minimalizuje hodnotu chyby ε . Rozvedme nyní naznačený postup poněkud podrobněji. Fourierův obraz impulzové charakteristiky filtru (tj. skutečnou frekvenční charakteristiku) zapíšeme ve tvaru

$$H(k, l) = \sum_{r=0}^{R-1} \sum_{s=0}^{S-1} h(r, s) \Phi_{k,l}(r, s), \quad (5.18)$$

kde

$$\Phi_{k,l}(r, s) = \exp \left[-j2\pi \left(\frac{rk}{M} + \frac{sl}{N} \right) \right]. \quad (5.19)$$

Dosazením vztahu (5.18) do (5.17) dostáváme

$$\varepsilon = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} W_{k,l} \left[\sum_{r=0}^{R-1} \sum_{s=0}^{S-1} h(r, s) \Phi_{k,l}(r, s) - I(k, l) \right]^2. \quad (5.20)$$

Pro nalezení extrému položíme

$$\frac{\partial \varepsilon}{\partial h(m, n)} = 0, \quad (m = 0, 1, \dots, R-1, n = 0, 1, \dots, S-1). \quad (5.21)$$

Po úpravách zjišťujeme, že hledané hodnoty $h(r, s)$ jsou řešením soustavy lineárních rovnic

$$\sum_{r=0}^{R-1} \sum_{s=0}^{S-1} A_{m,n}(r, s) h(r, s) = B_{m,n}, \quad (5.22)$$

kde

$$A_{m,n}(r, s) = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} W_{k,l} \Phi_{k,l}(r, s) \Phi_{k,l}(m, n),$$

$$B_{m,n} = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} W_{k,l} I(k, l) \Phi_{k,l}(m, n).$$

5.2 Rekurzivní filtrace

V případě rekurzivní filtrace je výstupní signál g svázán se vstupním signálem f prostřednictvím vztahu

$$b(m, n) * g(m, n) = a(m, n) * f(m, n), \quad (5.23)$$

kde b , a jsou diskrétní funkce, které rekurzivní filtr popisují. Nechť je hodnota $b(0, 0)$ nenulová. Beze ztráty na obecnosti předpokládáme, že $b(0, 0) = 1$. Rozepsáním konvoluce ze vztahu (5.23) snadno ověříme, že platí

$$g(m, n) = a(m, n) * f(m, n) + c(m, n) * g(m, n), \quad (5.24)$$

kde funkce c je definována takto

$$c(m, n) = \begin{cases} 0, & m = 0, n = 0 \\ -b(m, n), & \text{jinak} \end{cases}. \quad (5.25)$$

Fourierovou transformací rovnice (5.23) obdržíme vztah

$$G(k, l) = \frac{A(k, l)}{B(k, l)} F(k, l), \quad (5.26)$$

kde A , B , F , G jsou Fourierovy obrazy funkcí a , b , f , g . Vidíme, že frekvenční charakteristika rekurzivního filtru je $H(k, l) = A(k, l)/B(k, l)$.

Motivací pro použití rekurzivní filtrace je zpravidla to, že rekurzivní filtr může mít nižší časovou složitost než nerekurzivní filtr obdobných vlastností. Nechť N , N_a , N_b jsou rozměry oblastí, nad nimiž funkce f , a , b nabývají nenulových hodnot (pro jednoduchost nyní předpokládáme, že všechny oblasti mají čtvercový tvar). Z výrazu (5.24) je zřejmé, že časová složitost rekurzivní filtrace při přímém výpočtu je (detailněji je postup výpočtu popsán v následující podkapitole)

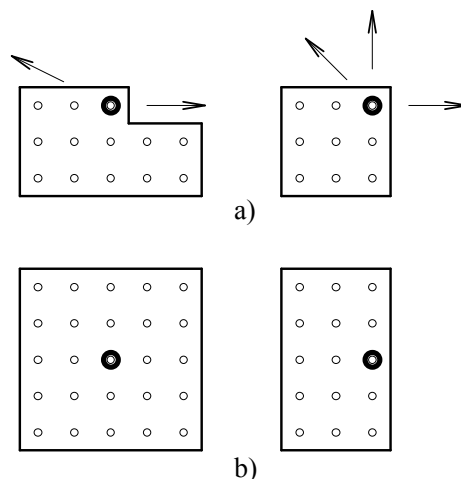
$$C_R = N^2 (N_a^2 + N_b^2). \quad (5.27)$$

Rekurzivní filtr je tedy výhodný tehdy, jestliže se funkce a , b podaří navrhnout tak, aby rozměry N_a , N_b byly malé (menší než rozměr oblasti, nad níž nabývá nenulových hodnot funkce h nerekurzivního filtru srovnatelných vlastností). Poznamenejme, že hodnoty N_a , N_b mohou být u rekurzivního filtru konečné i tehdy, jestliže je impulzová charakteristika filtru nenulová nad nekonečným počtem bodů.

5.2.1 Realizace rekurzivního filtru přímým výpočtem

Ze vztahu (5.24) vidíme, že k tomu, aby bylo možné provést výpočet hodnoty $g(m, n)$ výstupního pole, je nutné dříve znát ty hodnoty pole g , pro které jsou odpovídající hodnoty funkce c nenulové. Výpočet je proto žádoucí zorganizovat tak, aby hodnoty g byly získávány v potřebném pořadí. Je ovšem zřejmé, že pro některé funkce

c nebude existovat žádné pořadí, které by přímý výpočet umožnilo (později však uvidíme že, takové filtry mohou být vyčíslitelné iteračním postupem). Na obr. 5.2 jsou ilustrovány oba případy. Je zde vyznačen tvar oblasti, nad níž funkce c nabývá nenulových hodnot, a také možné směry postupu při výpočtu. Plným kruhem je vyznačeno místo, pro které je hodnota $g(m, n)$ výstupního pole v daném okamžiku určována.



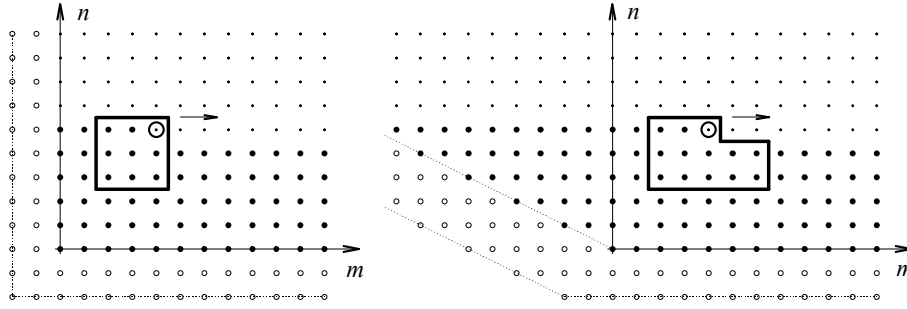
Obr. 5.2: a) Funkce $c(m, n)$ umožňující přímý výpočet, b) neumožňující přímý výpočet

Pro „nastartování“ výpočtu rekurzivního filtru je třeba mít předem k dispozici některé hodnoty výstupního obrazu g . Tyto hodnoty mají význam okrajových podmínek. Požadujeme-li filtr lineární, pak je nutné jako okrajové podmínky zadávat pouze hodnoty 0. Toto tvrzení vyplývá z jednoduché úvahy: Necht \mathcal{O} je uvažovaný rekurzivní filtr. Předpokládejme, že filtr \mathcal{O} je lineární. Platí proto

$$\mathcal{O} \{af(m, n)\} = a\mathcal{O} \{f(m, n)\}. \quad (5.28)$$

Položíme-li v (5.28) $a = 0$, konstatujeme na základě uvedeného vztahu, že v případě lineárního filtru nulovému vstupnímu obrazu nutně odpovídá také nulový obraz výstupní. Při volbě nenulových okrajových podmínek by zde obecně došlo ke sporu: nulovému vstupnímu obrazu by odpovídal nenulový obraz výstupní, a proto by filtr \mathcal{O} nebyl lineární. Ani poloha bodů, ve kterých zadáváme okrajové podmínky, nemůže být volena libovolně. Okrajové podmínky, lze zadávat pouze v těch místech, kde hodnotu výstupního pole neurčuje výraz (5.24). Obr. 5.3 ukazuje příklady míst, v nichž se zadávají okrajové podmínky, pro různé tvary oblasti, nad níž je funkce c nenulová.

Na rozdíl od nerekurzivních filtrů, které jsou vždy stabilní, je u filtrů rekurzivních nutno sledovat problémy stability. Jestliže je filtr nestabilní, pak se v průběhu



Obr. 5.3: Příklady okrajových podmínek pro různé funkce $c(m, n)$

výpočtu mohou šířit zaokrouhlovací chyby a případné šumy vstupního pole a dosahovat vysokých hodnot. Tento problém byl v minulosti detailně studován, ale jeho podrobnější diskuse by přesáhla rámec tohoto textu. V případě zájmu proto čtenáře odkazujeme např. na práci [9].

5.2.2 Realizace rekurzivního filtru iteračním výpočtem

V předchozím odstavci jsme vysvětlili, že má-li oblast, nad níž funkce b , c nabývají nenulové hodnoty, vhodný tvar, pak lze rekurzivní filtraci realizovat přímým výpočtem. Ne všechny filtry jsou ovšem takto realizovatelné. Přímý výpočet nelze uplatnit tehdy, jestliže by pro výpočet hodnoty $g(m, n)$ bylo zapotřebí hodnot v těch místech výstupního pole, kde zatím nebyly vypočteny. V těchto případech lze však použít iteračního postupu. Na základě vztahu (5.24) se pro výpočet $g(m, n)$ nabízí iterační předpis

$$g_i(m, n) = a(m, n) * f(m, n) + c(m, n) * g_{i-1}(m, n), \quad (5.29)$$

kde $g_i(m, n)$ znamená hodnotu v i -tém iteračním kroku. Pro zahájení iteračního pochodu volíme $g_0(m, n) = 0$. Oprávněnost iteračního postupu popsaného vztahem (5.29) je ovšem samozřejmě nutné doložit studiem konvergence. Ačkoli předpokládáme, že samotný výpočet filtrace bude prováděn v prostorové doméně, analýzu konvergence provedeme v doméně frekvenční. Fourierovou transformací vztahu (5.24) obdržíme

$$G(k, l) = A(k, l)F(k, l) + C(k, l)G(k, l), \quad (5.30)$$

kde C je Fourierovou transformací funkce c . Z (5.30) dostaneme dále vztah

$$G(k, l) = \frac{A(k, l)}{1 - C(k, l)} F(k, l). \quad (5.31)$$

Fourierovou transformací rovnice (5.29) získáme

$$G_i(k, l) = A(k, l)F(k, l) + C(k, l)G_{i-1}(k, l). \quad (5.32)$$

Odtud máme

$$G_1(k, l) = A(k, l)F(k, l), \quad (5.33)$$

$$G_2(k, l) = A(k, l)F(k, l) [1 + C(k, l)], \quad (5.34)$$

$$G_n(k, l) = A(k, l)F(k, l) \sum_{i=0}^{n-1} C^i(k, l).$$

S využitím identity

$$\sum_{i=0}^{n-1} C^i(k, l) = \frac{1 - C^n(k, l)}{1 - C(k, l)}, \quad (5.35)$$

za předpokladu, že $|C(k, l)| < 1$, a s uplatněním vztahu (5.31) docházíme v případě nekonečného počtu iteračních kroků k následujícímu závěru

$$\lim_{n \rightarrow \infty} G_n(k, l) = A(k, l)F(k, l) \lim_{n \rightarrow \infty} \frac{1 - C^n(k, l)}{1 - C(k, l)} = \frac{A(k, l)}{1 - C(k, l)} F(k, l) = G(k, l). \quad (5.36)$$

Vidíme tedy, že jestliže platí $|C(k, l)| < 1$, pak iterační proces konverguje k přesnému řešení. Poznamenejme však, že uvedená podmínka je zbytečně přísná. Podrobnější analýzou bylo totiž dokázáno, že iteračním postupem lze realizovat každý stabilní filtr [9].

5.2.3 Návrh rekurzivního filtru

Návrhem rekurzivního filtru rozumíme stanovení funkcí $a(m, n)$, $b(m, n)$ tak, aby měl filtr požadované vlastnosti. Východními údaji pro návrh může být např. nějaký vstupní signál $f(m, n)$ a požadovaná odezva $d(m, n)$ filtru na tento signál (příkladem je realizace filtru se zadanou impulzovou charakteristikou). Naznačme postup takového návrhu. Aby časová složitost filtrace byla co nejmenší, požadujeme, aby funkce a , b měly nenulové funkční hodnoty jen v malém počtu bodů. Necht $g(m, n)$ je skutečná odezva filtru na funkci $f(m, n)$. S ohledem na předchozí požadavek se bude skutečná odezva lišit od odezvy požadované. Úhrnná kvadratická chyba filtru v prostorové doméně pak je

$$\varepsilon = \sum_m \sum_n [g(m, n) - d(m, n)]^2. \quad (5.37)$$

Funkce $a(m, n)$, $b(m, n)$ stanovíme minimalizací funkcionálu (5.37). Analogický postup můžeme realizovat také ve frekvenční doméně. Úhrnná kvadratická chyba filtru ve frekvenční doméně je

$$\varepsilon = \sum_k \sum_l [G(k, l) - D(k, l)]^2 = \sum_k \sum_l \left[\frac{A(k, l)}{B(k, l)} F(k, l) - D(k, l) \right]^2. \quad (5.38)$$

Můžeme také zavést funkci $W_{k,l}$ vyjadřující váhu chyby pro jednotlivé frekvence. Pro chybu pak máme vztah

$$\varepsilon_W = \sum_k \sum_l W_{k,l} \left[\frac{A(k,l)}{B(k,l)} F(k,l) - D(k,l) \right]^2. \quad (5.39)$$

Poznamenejme na závěr, že ať již použijeme návrhu v doméně prostorové nebo frekvenční, jsou rezultující úlohy minimalizace nelineární, a proto je zde zapotřebí použít numerického řešení.

5.3 Obecný model degradace a rekonstrukce obrazu

Při snímání obrazu často dochází k jeho různým poškozením (degradacím). Obraz může být např. zašumělý, neostřý atd. Pokud je to možné, bývá obvykle žádoucí tyto degradace korigovat. Korekce bývá dosti často realizována filtrací. Klíčovou otázkou je volba vhodného filtru. K jeho nalezení je nejprve vhodné zavést model degradace obrazu. Označme f_I vstupní nepoškozený obraz. Obraz f_I není ovšem k dispozici. Dostupný je místo něj pouze obraz poškozený, který označíme f_D . Představujeme si, že degradace je způsobena nějakým operátorem (označme jej \mathcal{D}) a dále si představujeme, že dochází k degradaci aditivním šumem (označme jej v). Je tedy

$$f_D = \mathcal{D} \{f_I\} + v. \quad (5.40)$$

Úkolem rekonstrukce obrazu je nalézt rekonstrukční operátor \mathcal{R} tak, abychom jeho aplikací na obraz f_D obdrželi obraz (označme jej f_R), který je co možná nejbližší obrazu f_I . Máme tedy

$$f_R = \mathcal{R} \{f_D\} = \mathcal{R} \{ \mathcal{D} \{f_I\} + v \}. \quad (5.41)$$

Ze všech možných operátorů zvolíme za operátor \mathcal{R} ten, který dává minimální hodnotu chyby

$$\varepsilon = \|f_R - f_I\|. \quad (5.42)$$

Předpokládejme nyní, že degradační operátor \mathcal{D} je lineární a invariantní vůči posuvu. Jeho impulzovou a frekvenční charakteristiku označme h_D , H_D . Také rekonstrukční operátor \mathcal{R} předpokládáme lineární a invariantní vůči posuvu. Úkolem je nalézt impulzovou charakteristiku h_R nebo frekvenční charakteristiku H_R rekonstrukčního operátoru \mathcal{R} tak, aby byla minimalizována chyba ze vztahu (5.42). Zapišme rovnice (5.40), (5.41) pomocí konvoluce. Máme

$$f_D(m, n) = f_I(m, n) * h_D(m, n) + v(m, n), \quad (5.43)$$

$$f_R(m, n) = [f_I(m, n) * h_D(m, n) + v(m, n)] * h_R(m, n). \quad (5.44)$$

Provedeme-li Fourierovu transformaci uvedených vztahů, dostaneme

$$F_D(k, l) = F_I(k, l)H_D(k, l) + V(k, l). \quad (5.45)$$

$$F_R(k, l) = [F_I(k, l)H_D(k, l) + V(k, l)] H_R(k, l). \quad (5.46)$$

5.4 Inverzní filtrace

O *inverzním filtru* hovoříme tehdy, jestliže frekvenční charakteristiku rekonstrukčního operátoru \mathcal{R} zvolíme ve tvaru

$$H_R(k, l) = \frac{1}{H_D(k, l)}. \quad (5.47)$$

Dosazením vztahu (5.47) do (5.46) zjišťujeme, že platí

$$F_R(k, l) = F_I(k, l) + \frac{V(k, l)}{H_D(k, l)}. \quad (5.48)$$

Inverzní Fourierovou transformací pak obdržíme

$$f_R(m, n) = f_I(m, n) + \mathcal{F}^{-1} \left\{ \frac{V(k, l)}{H_D(k, l)} \right\}. \quad (5.49)$$

Ze vztahu (5.49) vidíme, že pokud by obraz f_D nebyl degradován šumem, pak by inverzní filtr poskytl obraz, který je shodný s nepoškozeným vstupním obrazem f_I . Za přítomnosti šumu se ovšem obrazy f_R a f_I liší. Odchylka je popsána druhým členem na pravé straně rovnice (5.49). Je zřejmé, že vliv šumu roste s tím, jak klesají hodnoty $H_D(k, l)$. Protože u většiny degradací jsou hodnoty $H_D(k, l)$ odpovídající vyšším frekvenčním složkám, kde vliv šumu zejména očekáváme, malé (obraz je obvykle rozostřován tím, že jsou potlačeny vyšší frekvence), dochází použitím inverzního filtru často k nežádoucímu zdůraznění šumu. Jiným praktickým problémem je, že frekvenční charakteristiku degradačního operátoru mnohdy neznáme přesně. Často je nutné spokojit se s pouhým odhadem.

5.5 Wienerův filtr ve frekvenční doméně

V této podkapitole využijeme k nalezení rekonstrukčního filtru stochastického přístupu. Ukážeme odvození tzv. *Wienerova filtru* ve frekvenční doméně. Předpokládáme, že $\mathbf{f}_I(m, n)$, $\mathbf{f}_D(m, n)$, $\mathbf{f}_R(m, n)$ jsou homogenní náhodná pole rozměru $M \times N$ reprezentující postupně ideální, degradovaný a rekonstruovaný obraz a že $\mathbf{v}(m, n)$ je náhodné pole rozměru $M \times N$, které reprezentuje aditivní šum. Také odpovídající Fourierovy obrazy $\mathbf{F}_I(m, n)$, $\mathbf{F}_D(m, n)$, $\mathbf{F}_R(m, n)$, $\mathbf{V}(m, n)$ jsou náhodná pole rozměru $M \times N$. Pro zestručnění položíme

$$\mathbf{f}_\Delta(m, n) = \mathbf{f}_R(m, n) - \mathbf{f}_I(m, n), \quad (5.50)$$

$$\mathbf{F}_\Delta(k, l) = \mathcal{F} \{ \mathbf{f}_\Delta(m, n) \} = \mathbf{F}_R(k, l) - \mathbf{F}_I(k, l). \quad (5.51)$$

Rekonstrukční filtr nalezneme minimalizací chyby

$$\varepsilon = \mathbb{E} \left\{ \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |\mathbf{f}_\Delta(m, n)|^2 \right\}. \quad (5.52)$$

Podle Parsevalova teorému (2.63) ovšem platí

$$\varepsilon = \mathbb{E} \left\{ \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |\mathbf{f}_\Delta(m, n)|^2 \right\} = \mathbb{E} \left\{ \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} |\mathbf{F}_\Delta(k, l)|^2 \right\}. \quad (5.53)$$

$\mathbf{F}_\Delta(k, l)$ vyjádříme dosazením vztahu (5.46) do vztahu (5.51). Dostaneme

$$\mathbf{F}_\Delta(k, l) = [\mathbf{F}_I(k, l)H_D(k, l) + \mathbf{V}(k, l)] H_R(k, l) - \mathbf{F}_I(k, l). \quad (5.54)$$

Uvážíme-li, že $|\mathbf{F}_\Delta(k, l)|^2 = \mathbf{F}_\Delta(k, l)\mathbf{F}_\Delta^*(k, l)$, pak ze vztahu (5.53) máme

$$\varepsilon = \mathbb{E} \left\{ \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} \mathbf{F}_\Delta(k, l)\mathbf{F}_\Delta^*(k, l) \right\}. \quad (5.55)$$

Frekvenční charakteristiku H_R rekonstrukčního filtru určíme z podmínky

$$\frac{\partial \varepsilon}{\partial H_R(r, s)} = 0, \quad r = 0, 1, \dots, M-1, \quad s = 0, 1, \dots, N-1. \quad (5.56)$$

Ze vztahu (5.55) máme

$$\begin{aligned} \frac{\partial \varepsilon}{\partial H_R(r, s)} &= \mathbb{E} \left\{ \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} \left[\frac{\partial \mathbf{F}_\Delta(k, l)}{\partial H_R(r, s)} \mathbf{F}_\Delta^*(k, l) + \mathbf{F}_\Delta(k, l) \frac{\partial \mathbf{F}_\Delta^*(k, l)}{\partial H_R(r, s)} \right] \right\} \\ &= \mathbb{E} \left\{ \frac{\partial \mathbf{F}_\Delta(r, s)}{\partial H_R(r, s)} \mathbf{F}_\Delta^*(r, s) + \mathbf{F}_\Delta(r, s) \frac{\partial \mathbf{F}_\Delta^*(r, s)}{\partial H_R(r, s)} \right\}. \end{aligned} \quad (5.57)$$

Předpokládejme, že náhodná pole \mathbf{f}_I , \mathbf{v} jsou nekorelovaná. Uplatněním výsledku z podkapitoly 3.6 lze snadno ukázat, že pole $\mathbf{F}_I(k, l)$ a $\mathbf{V}(k, l)$ jsou proto také nekorelovaná, z čehož vyplývá $\mathbb{E}\{\mathbf{F}_I(k, l)\mathbf{V}(k, l)\} = \mathbb{E}\{\mathbf{F}_I(k, l)\}\mathbb{E}\{\mathbf{V}(k, l)\}$. Dále předpokládejme, že $\mathbb{E}\{\mathbf{v}(m, n)\} = 0$. Odtud plyne, že $\mathbb{E}\{\mathbf{V}(k, l)\} = 0$. Dosadíme-li do vztahu (5.57) vztah (5.54), položíme-li $\mathbb{E}\{\mathbf{F}_I(k, l)\mathbf{F}_I^*(k, l)\}/\sqrt{(MN)} = G_{\mathbf{f}_I\mathbf{f}_I}(k, l)$, $\mathbb{E}\{\mathbf{V}(k, l)\mathbf{V}^*(k, l)\}/\sqrt{(MN)} = G_{\mathbf{v}\mathbf{v}}(k, l)$, pak po úpravách dostaneme pro rekonstrukční filtr předpis

$$H_R(r, s) = H_D^*(r, s) \frac{G_{\mathbf{f}_I\mathbf{f}_I}(r, s)}{|H_D(r, s)|^2 G_{\mathbf{f}_I\mathbf{f}_I}(r, s) + G_{\mathbf{v}\mathbf{v}}(r, s)}. \quad (5.58)$$

Poznamenejme, že $G_{\mathbf{f}_I\mathbf{f}_I}$ je výkonová spektrální hustota vstupního nedegradovaného signálu a $G_{\mathbf{v}\mathbf{v}}$ je výkonová spektrální hustota šumu (viz. výsledek příkladu 3.1).

Kapitola 6

Bodové, algebraické a geometrické operace s obrazy

V praxi se často uplatňují operace, při nichž se modifikuje jas (případně intenzita barevných složek) tak, že jas výsledného obrazu v každém bodě je nějakou funkcí jasu původního obrazu v témže bodě. Takové operace nazýváme operacemi bodovými. Do operací algebraických vstupují zpravidla dva obrazy. Výsledný obraz se získává tak, že se jasy ve stejnohlých bodech vstupních obrazů sčítají, odčítají, násobí nebo dělí. Na rozdíl od bodových a algebraických operací mění operace geometrické tvar a rozměry obrazu. O uvedených třídách operací pojednáme v této kapitole podrobněji.

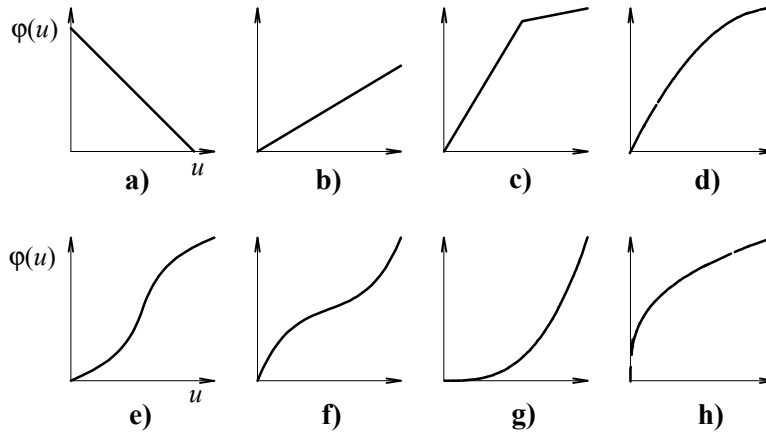
6.1 Bodové operace

Nechť $f(x, y)$ je obrazová funkce (jasová funkce) popisující vstupní obraz a $g(x, y)$ obrazová funkce popisující obraz výstupní. Bodovou operací rozumíme transformaci obrazu podle předpisu

$$g(x, y) = \varphi[f(x, y)]. \quad (6.1)$$

Ze vztahu (6.1) je zřejmé, že bodové operace vytvářejí výstupní obraz tak, že bodu o souřadnicích (x, y) přiřadí hodnotu jasu, která je nějakou funkcí φ jasu vstupního obrazu v bodě o stejných souřadnicích. Důvody pro provádění bodových transformací mohou být rozmanité. Uvedme několik příkladů: Citlivost prvků, které se při snímání obrazu používají často není lineární. Při fotometrické kalibraci kompenzujeme bodovou operací nelinearitu snímacího prvku. Podobná je i situace na výstupu řetězce zpracovávajícího obraz, kde bodovou operací můžeme kompenzovat nelinearitu displeje. V systémech pro zpracování obrazu je jednou z nejpoužívanějších operací bodová operace provádějící změnu rozsahu nebo průběhu jasu.

Předpokládáme, že jas (označme jej u) vstupního obrazu může nabývat hodnot z intervalu $\langle u_{\min}, u_{\max} \rangle$ a požadujeme, aby hodnoty u_{\min}, u_{\max} byly transformovány



Obr. 6.1: Příklady funkcí realizujících bodové operace: **a)**, **b)** lineární funkce dle vztahu (6.3), **c)** funkce po částech lineární, **d)** kvadratická parabola (6.4) ($c = 1$), **e)**, **f)** kubická parabola (6.5), **g)**, **h)** exponenciální funkce (6.6) ($\gamma = 3$, $\gamma = 1/3$)

na hodnoty $v_1 = \varphi(u_{\min})$, $v_2 = \varphi(u_{\max})$, které považujeme za známé. Zavedme hodnotu ξ vztahem

$$\xi = \frac{u - u_{\min}}{u_{\max} - u_{\min}}. \quad (6.2)$$

Jako příklady funkcí $\varphi(\xi)$ realizujících bodovou transformaci lze uvést např. následující funkce:

$$\varphi(\xi) = v_1 + \xi(v_2 - v_1), \quad (6.3)$$

$$\varphi(\xi) = v_1 + [\xi + c\xi(1 - \xi)](v_2 - v_1), \quad (6.4)$$

$$\varphi(\xi) = (1 - 3\xi^2 + 2\xi^3)v_1 + (\xi - 2\xi^2 + \xi^3)v_1' + (3\xi^2 - 2\xi^3)v_2 + (-\xi^2 + \xi^3)v_2', \quad (6.5)$$

$$\varphi(\xi) = v_1 + \xi^\gamma(v_2 - v_1). \quad (6.6)$$

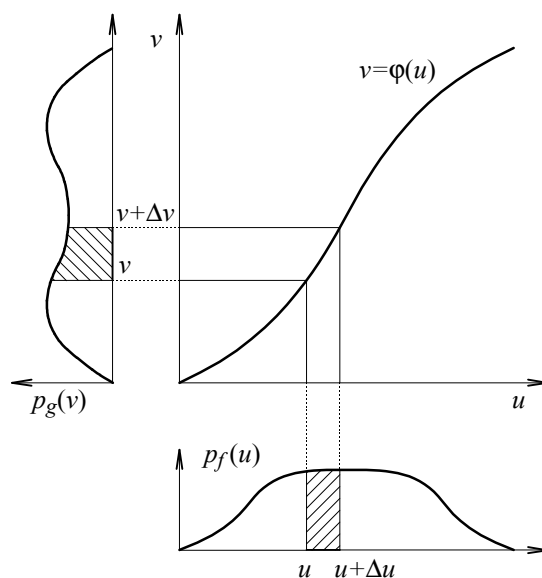
kde $v_1' = \varphi'(u)|_{u=u_{\min}}$, $v_2' = \varphi'(u)|_{u=u_{\max}}$ a c, γ jsou konstanty zvolené tak, aby bylo dosaženo žádaného průběhu funkce φ . Příklady průběhů funkcí jsou uvedeny na obr 6.1.

6.1.1 Vliv bodových operací na histogram jasu

Nejprve uvedeme obecné úvahy týkající se *histogramu jasu*. Uvažujme diskrétní obrazy s jasem popsaným celými čísly. Obraz je definován nad oblastí sestávající

z N bodů. Necht N_z označuje počet bodů, jejichž jas je právě z . Pro histogram (označme jej H) jasu pak máme $H(z) = N_z$. Probírejme postupně jednotlivé body obrazu a vyšetřujme jejich jas. Na hodnotu jasu můžeme pohlížet jako na náhodnou proměnnou (označme ji \mathbf{b}). Hodnota $p(z) = H(z)/N = N_z/N$ udává pravděpodobnost jevu, že jas bodu je právě z . Je zřejmé, že $H(z) = Np(z)$. Poznamenejme, že pro hodnotu $p(z)$ bývá též někdy používáno názvu normalizovaný histogram. Pro teoretické úvahy je výhodné zavést odpovídající pojem také pro spojité obrazy s reálnými hodnotami jasu. Necht $\mathcal{P}\{z < \mathbf{b} \leq z + \Delta z\}$ označuje pravděpodobnost jevu, že jas bodu nabývá hodnoty z intervalu $(z, z + \Delta z)$, \mathcal{S} necht je plocha celého obrazu a $\mathcal{S}\{z < \mathbf{b} \leq z + \Delta z\}$ necht je plocha těch jeho částí, kde jas nabývá hodnoty z intervalu $(z, z + \Delta z)$. Ve spojitém případě má $p(z)$ význam hustoty pravděpodobnosti

$$p(z) = \lim_{\Delta z \rightarrow 0} \frac{\mathcal{P}\{z < \mathbf{b} \leq z + \Delta z\}}{\Delta z} = \lim_{\Delta z \rightarrow 0} \frac{\mathcal{S}\{z < \mathbf{b} \leq z + \Delta z\}}{\mathcal{S} \Delta z} \quad (6.7)$$



Obr. 6.2: Vliv bodových operací na histogram obrazu

Vyšetřujme nyní, jak se histogram jasu mění vlivem bodových operací. Uvažujme spojité obrazy s reálnými hodnotami jasu. Předpokládejme, že hustota pravděpodobnosti jasu vstupního obrazu f je známa (označme ji p_f) a že na obraz f byla aplikována bodová operace popsaná rovnicí (6.1). Ukážeme, jak lze stanovit hustotu pravděpodobnosti jasu výstupního obrazu g (označme ji p_g). Necht u, v značí vstupní resp. výstupní jas. Je tedy $v = \varphi(u)$. Předpokládejme, že φ je bijekcí, a že lze proto konstruovat inverzní funkci $u = \varphi^{-1}(v)$. Uvažujme interval $\langle u, u + \Delta u \rangle$ vstupních jasů. Tomuto intervalu odpovídá interval $\langle v, v + \Delta v \rangle = \langle \varphi(u), \varphi(u + \Delta u) \rangle$ jasů výstupních (obr. 6.2). Řekněme, že ve vstupním obraze měla jistá plocha obrazu

jas z intervalu $\langle u, u + \Delta u \rangle$. Po transformaci bude mít tatáž plocha jas z intervalu $\langle v, v + \Delta v \rangle$. Vzhledem ke vztahu (6.7) proto máme

$$\int_v^{v+\Delta v} p_g(z) dz = \int_u^{u+\Delta u} p_f(z) dz. \quad (6.8)$$

Jestliže jsou hodnoty Δu , Δv dostatečně malé, pak můžeme na základě vztahu (6.8) také psát

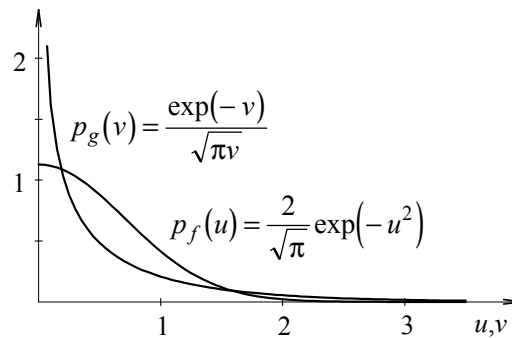
$$p_g(v)\Delta v = p_f(u)\Delta u. \quad (6.9)$$

Odtud získáme pro výpočet hustoty pravděpodobnosti $p_g(v)$ předpis

$$p_g(v) = \frac{p_f(u)}{\Delta v / \Delta u} = \frac{p_f(u)}{\varphi'(u)} = \frac{p_f[\varphi^{-1}(v)]}{\varphi'[\varphi^{-1}(v)]}. \quad (6.10)$$

Příklad 6.1. Předpokládejme, že jasy vstupního obrazu jsou nezáporné ($u \geq 0$) a že hustota pravděpodobnosti jasu vstupního obrazu je popsána výrazem

$$p_f(u) = \frac{2}{\sqrt{\pi}} \exp(-u^2).$$



Obr. 6.3: Hustoty pravděpodobností jasu z příkladu 6.1

Výstupní obraz je definován vztahem

$$g(x, y) = f^2(x, y).$$

Určíme hustotu pravděpodobnosti jasu výstupního obrazu. Máme $v = \varphi(u) = u^2$, $u = \varphi^{-1}(v) = \sqrt{v}$. Podle vztahu (6.10) pak pro hledanou hustotu dostáváme

$$p_g(v) = \frac{\exp(-v)}{\sqrt{\pi v}}.$$

6.1.2 Vyrovnání histogramu

Vyrovnáním histogramu rozumíme provedení takové bodové operace, která zajistí rovnoměrné rozložení hustoty pravděpodobnosti jasu výsledného obrazu. Důvody k provedení operace mohou být následující: 1) Vyrovnání může vést ke zlepšení kvality obrazu subjektivně vnímané pozorovatelem. 2) Vyrovnání může být použito jako transformace zajišťující normalizované jednotné podmínky před dalším zpracováním obrazů (např. před segmentací a rozpoznáním). (Poznamenejme, že pro jistou třídu obrazů může však vyrovnání histogramu vést naopak i ke zhoršení subjektivně vnímané kvality. Jako příklad uvažte např. obraz vzniklý naskenováním výkresu obsahujícího černé čáry na bílém podkladě.) Zaměříme se nyní na problém stanovení funkce φ ze vztahu (6.1) tak, aby rozložení pravděpodobnosti p_g bylo rovnoměrné. Úvahy se zjednoduší, jestliže budeme předpokládat, že jas vstupního i výstupního obrazu je reprezentován reálnými čísly (obr. 6.4a). Z podmínky rovnoměrného rozložení hustoty pravděpodobnosti máme $p_g(v) = 1/(v_{\max} - v_{\min})$. Použitím vztahu (6.10) pak dostaneme

$$\varphi'(v) = \frac{p_f(u)}{p_g(v)} = p_f(u)(v_{\max} - v_{\min}). \quad (6.11)$$

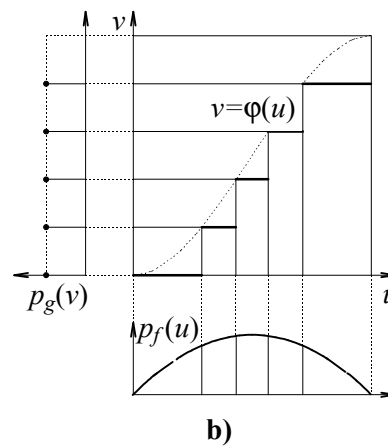
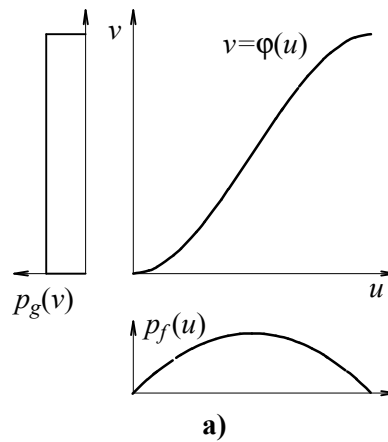
Odtud

$$\varphi(u) = \int_{u_{\min}}^u \varphi'(z) dz = (v_{\max} - v_{\min}) \int_{u_{\min}}^u p_f(z) dz. \quad (6.12)$$

Protože je podle předpokladu hustota $p_f(z)$ známá, je vztah (6.12) předpisem pro výpočet hledané funkce φ . Jestliže jsou jasy reprezentovány celými čísly, je zapotřebí několik dalších upřesnění. Zaměříme se nejprve na případ, kdy je celými čísly reprezentován pouze jas výstupního obrazu. Máme tedy $v \in \{v_{\min}, v_{\min} + 1, \dots, v_{\max}\}$. Jasy vstupního obrazu jsou reálné. Funkce φ má v tomto případě stupňovitý tvar (obr. 6.4b). Pro $z \in (u_v, u_{v+1})$ platí $\varphi(z) = v$. Problémem je stanovit pro jednotlivé hodnoty v odpovídající hodnoty u_v . S přihlédnutím ke skutečnosti, že v tomto případě je $p_g(v) = 1/(v_{\max} - v_{\min} + 1)$, dostáváme z podmínky rovnoměrného rozdělení pravděpodobností pro hodnotu $u_v (v_{\min} \leq v \leq v_{\max})$ následující rovnici

$$\frac{v - v_{\min}}{(v_{\max} - v_{\min} + 1)} = \int_{u_{\min}}^{u_v} p_f(z) dz. \quad (6.13)$$

Jestliže je celými čísly reprezentován jas jak výstupního tak i vstupního obrazu, pak je funkce φ diskrétní. Stejně jako v předchozím případě, je i zde problémem stanovit pro jednotlivé hodnoty v odpovídající hodnoty u_v . Hledanou hodnotu u_v lze určit jako nejmenší hodnotu, pro kterou je splněna nerovnost



Obr. 6.4: Vyrovnání histogramu jasu

$$\frac{v - v_{\min}}{(v_{\max} - v_{\min} + 1)} \leq \sum_{z=u_{\min}}^{u_v} p_f(z). \quad (6.14)$$

Jak z uvedeného postupu vyplývá, nelze v případě, kdy je celými čísly reprezentován jas výstupního i vstupního obrazu, dosáhnout vyrovnání histogramu zcela přesně. Histogram je vyrovnán pouze přibližně.

6.2 Algebraické operace s obrazy

Nechť $f_1(x, y)$, $f_2(x, y)$ jsou známé obrazy. Algebraickou operací rozumíme postup, kdy je výsledný obraz $g(x, y)$ získáván pomocí některého z následujících vztahů:

$$g(x, y) = f_1(x, y) + f_2(x, y), \quad (6.15)$$

$$g(x, y) = f_1(x, y) - f_2(x, y), \quad (6.16)$$

$$g(x, y) = f_1(x, y) \times f_2(x, y), \quad (6.17)$$

$$g(x, y) = f_1(x, y) \div f_2(x, y). \quad (6.18)$$

Uvedme několik příkladů, kdy mohou být algebraické operace použity v praxi: Při snímání kamerou je někdy vlivem nedokonalosti objektivu obrázků uprostřed světlejší a na krajích (nejvíce v rozích) tmavší. Uvedenou vadu objektivu můžeme považovat za multiplikační chybu a korigovat ji pomocí vztahů (6.17) nebo (6.18). Obraz f_1 je obraz, který má být korigován. Korekční hodnoty obsažené v obrazu f_2 snadno vypočítáme např. z hodnot získaných sejmutím bílé rovnoměrně osvětlené plochy. Podobně lze korigovat systematickou chybu vznikající nerovnoměrným osvětlením předlohy při snímání kamerou. Při snímání scény, v níž se pohybují objekty, lze rozdíl po sobě jdoucích obrazů využít k detekci pohybu. Použití sčítání obrazu pro redukci šumu ukážeme podrobněji v samostatné podkapitole.

6.2.1 Vliv algebraických operací na histogram jasu

Nechť $p_1(z)$, $p_2(z)$ jsou hustoty pravděpodobnosti jasu v obrazech $f_1(x, y)$, $f_2(x, y)$. Stanovíme hustotu pravděpodobnosti $p(z)$ jasu obrazu $g(x, y)$ vzniklého součtem dle vztahu (6.15). Označme $p_{1,2}(u, v)$ sdruženou hustotu pravděpodobnosti jasu v prvním a druhém obraze. Předpokládáme, že vstupní obrazy jsou nezávislé, a že proto platí $p_{1,2}(u, v) = p_1(u)p_2(v)$. Podle vztahu (6.15) je hodnoty z jasu v obraze $g(x, y)$ dosaženo tehdy, jestliže při jasu v v obraze f_2 je jas v obraze f_1 roven hodnotě $u = z - v$. Odtud vychází pro hledanou hustotu $p(z)$ vztah

$$p(z) = \int_{-\infty}^{\infty} p_{1,2}(z - v, v) dv = \int_{-\infty}^{\infty} p_1(z - v)p_2(v) dv = p_1(z) * p_2(z). \quad (6.19)$$

Vztah (6.19) ukazuje, že hustota pravděpodobnosti jasu výsledného obrazu je rovna konvoluci hustot pravděpodobností jasů obou obrazů vstupních. Získaný výsledek lze snadno rozšířit také na rozdíl obrazů dle vztahu (6.16). Postačí položit $\hat{f}_2(x, y) = -f_2(x, y)$, čímž převedeme rozdíl na součet. Pro hustotu pravděpodobnosti jasu obrazu \hat{f}_2 máme $\hat{p}_2(z) = p_2(-z)$.

6.2.2 Využití průměrování k redukci šumu

V tomto odstavci podrobněji popíšeme jednu z praktických aplikací sčítání obrazů. Nechť $f(x, y)$ je teoretický obraz neměnný v čase. Při snímání získáváme ovšem obraz, který je poškozen šumem. Obraz proto snímáme opakovaně M -krát. Při i -tém

sejmutí obdržíme obraz $f(x, y) + n_i(x, y)$, kde n_i je šum. Za aproximaci teoretického obrazu můžeme intuitivně brát průměr

$$\hat{f}(x, y) = \frac{1}{M} \sum_{i=1}^M [f(x, y) + n_i(x, y)]. \quad (6.20)$$

Ukážeme, že použití průměru podle rovnice (6.20) objektivně vede ke zlepšení kvality sejmutého obrazu. K hodnocení kvality použijeme poměru signál-šum, který je definován vztahem

$$SNR(x, y) = \frac{f(x, y)}{\sqrt{E\{n^2(x, y)\}}}. \quad (6.21)$$

Pro hodnotu poměru signál-šum v obraze vzniklém průměrováním M obrazů odtud vychází

$$SNR_M(x, y) = \frac{f(x, y)}{\sqrt{E\left\{\left[\frac{1}{M} \sum_{i=1}^M n_i(x, y)\right]^2\right\}}}. \quad (6.22)$$

Výraz (6.22) dále upravíme. Při tom využijeme následujících předpokladů: 1) $n_i(x, y)$ je náhodná proměnná, pro níž platí $E\{n_i(x, y)\} = 0$; 2) náhodné proměnné $n_i(x, y)$, $n_j(x, y)$ jsou nekorelované, a proto platí $E\{n_i(x, y)n_j(x, y)\} = E\{n_i(x, y)\}E\{n_j(x, y)\}$; 3) $E\{n_i^2(x, y)\} = E\{n_j^2(x, y)\} = E\{n^2(x, y)\}$. Postupně dostáváme

$$SNR_M(x, y) = \frac{Mf(x, y)}{\sqrt{E\left\{\sum_{i=1}^M n_i^2(x, y)\right\}}} = \sqrt{M} \frac{f(x, y)}{\sqrt{E\{n^2(x, y)\}}} = \sqrt{M} SNR(x, y). \quad (6.23)$$

Vztah (6.23) ukazuje, že průměrováním z M obrazů se poměr signál-šum zvětšil \sqrt{M} -krát.

6.3 Geometrické transformace obrazu

Geometrické transformace mění tvar nebo rozměry obrazu. Obecně jsou popsány vztahem

$$g(x, y) = f[\varphi(x, y), \psi(x, y)]. \quad (6.24)$$

Je zřejmé, že geometrická transformace je zcela zadána funkcemi $\varphi(x, y)$, $\psi(x, y)$. Vztah (6.24) navíc podává i praktický návod na to, jak transformaci provést. Předpokládejme, že obrazy $f(x, y)$ i $g(x, y)$ jsou diskrétní. Na základě vztahu (6.24) lze postupně probírat všechny body výstupního obrazu g a pro každý bod (x, y)

určit odpovídající polohu $[\varphi(x, y), \psi(x, y)]$ ve vstupním obraze f . Hodnota jasu $f[\varphi(x, y), \psi(x, y)]$ je pak vzata jako hodnota $g(x, y)$. Zatímco x, y nabývají celočíselných hodnot (postupujeme po celočíselných souřadnicích pixelů výstupního obrazu), jsou hodnoty $\varphi(x, y), \psi(x, y)$ obvykle reálné, nikoli celočíselné. Podle předpokladu je však i vstupní obraz f diskrétní, a proto i pro něj jsou hodnoty jasu známy jen v bodech o celočíselných souřadnicích. Pro získání hodnot v mezilehlých bodech o neceločíselných souřadnicích lze využít metod, které byly podrobněji popsány v podkapitole 4.4. V praxi se nejčastěji používá bilineární interpolace nebo interpolace vyššího řádu.

Zabývejme se nyní podrobněji funkcemi $\varphi(x, y), \psi(x, y)$. Nejobecnějším požadavkem je, aby dvojice $[\varphi(x, y), \psi(x, y)]$ byla homeomorfismem (nesplnění tohoto požadavku by způsobilo „roztrhání“ obrazu). Kromě toho bývají obvykle požadovány i jiné vlastnosti. Jsou-li funkce $\varphi(x, y), \psi(x, y)$ lineární, pak je transformace (6.24) transformací afinní. *Afinní transformace* je zadána šesti hodnotami $a_{11}, a_{12}, a_{21}, a_{22}, b_1, b_2$ a lze ji zapsat následujícím vztahem

$$\begin{bmatrix} \varphi(x, y) \\ \psi(x, y) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \quad (6.25)$$

Příklad 6.2. Uvažujme diskrétní obraz rozměru $M \times N$ bodů popsaný obrazovou funkcí $f(m, n)$ (obr. 6.5). Provedeme rotaci obrazu kolem levého dolního rohu o úhel φ ($0 \leq \alpha \leq \pi/2$). Rozměry obrazu vzniklého rotací budou $[(N-1)\sin\varphi + (M-1)\cos\varphi + 1] \times [(N-1)\cos\varphi + (M-1)\sin\varphi + 1]$ bodů. Počátek $(0, 0)$ předpokládáme u obou obrazů v levém dolním rohu. Označme \tilde{x}, \tilde{y} souřadnice v původním a x, y souřadnice v novém obraze. Pro transformaci souřadnic z původního do nového obrazu snadno získáme předpis

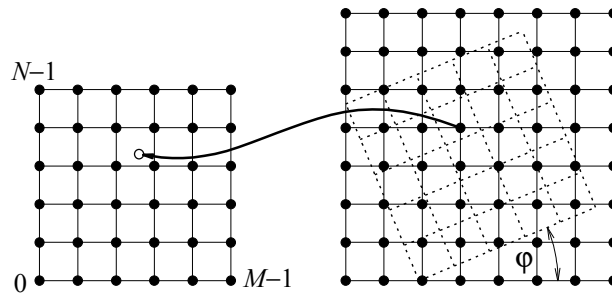
$$x = \tilde{x} \cos \varphi - \tilde{y} \sin \varphi + (N - 1) \sin \varphi,$$

$$y = \tilde{x} \sin \varphi + \tilde{y} \cos \varphi.$$

Odtud máme

$$\begin{bmatrix} \varphi(x, y) \\ \psi(x, y) \end{bmatrix} = \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = \begin{bmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + (N - 1) \begin{bmatrix} -\sin \varphi \cos \varphi \\ \sin^2 \varphi \end{bmatrix}.$$

Prakticky provedeme transformaci tak, že probíráme všechny body výstupního obrazu jeden po druhém. Poloha bodu ve výstupním obraze je specifikována celočíselnou dvojicí x, y . Jestliže platí $[0 \leq \varphi(x, y) \leq M - 1] \& [0 \leq \psi(x, y) \leq N - 1]$, určíme hodnotu $g(x, y)$ jako hodnotu $f[\varphi(x, y), \psi(x, y)]$, kterou vypočítáme interpolací. V opačném případě položíme např. $g(x, y) = 0$.



Obr. 6.5: Afinní transformace obrazu - rotace

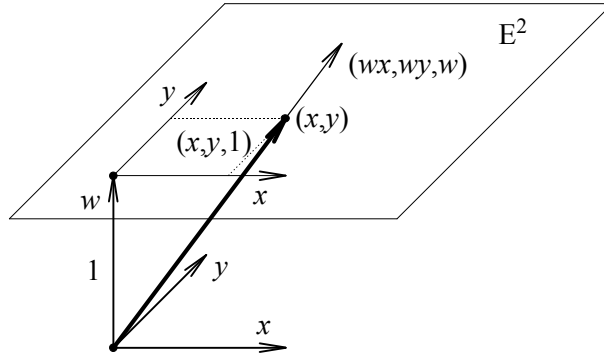
Příklad 6.3. Geometrickou transformací, která se v praxi uplatňuje velmi často, je změna rozměru obrazu. Opět uvažujme diskrétní obraz rozměru $M \times N$ bodů. Tento obraz má být transformován tak, aby rozměry výstupního obrazu byly $P \times Q$. Snadno obdržíme vztah

$$\begin{bmatrix} \varphi(x, y) \\ \psi(x, y) \end{bmatrix} = \begin{bmatrix} \frac{P-1}{M-1} & 0 \\ 0 & \frac{Q-1}{N-1} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Obecnější než transformace afinní jsou transformace projektivní. Jak afinní tak projektivní transformace zachovávají linearitu transformovaných útvarů (přímky se transformují opět na přímky). Na rozdíl od afinních transformací však projektivní transformace nezachovávají rovnoběžnost. Projektivní transformace se uplatňují zejména v případech, které nějakým způsobem souvisí se středovým promítáním. Příkladem praktické úlohy je transformace, při níž má být obraz deformován tak, aby budil dojem, že je umístěn na nějaké obecně umístěné rovině v prostoru, kterou zobrazujeme středovým promítáním.

Při popisu projektivních transformací se zpravidla používá homogenních souřadnic [17]. *Homogenní souřadnice* reprezentují bod v n -rozměrném afinním prostoru jako směr v jistém přidruženém $(n+1)$ -rozměrném vektorovém prostoru. Uvažujme dvojrozměrný afinní prostor a zavedme přidružený trojrozměrný vektorový prostor a souřadnou soustavu v něm tak, jak je znázorněno na obr. 6.6. V afinním prostoru uvažujme bod, jehož afinní souřadnice jsou (x, y) . Homogenními souřadnicemi uvažovaného bodu je trojice (wx, wy, w) , kde w je libovolné reálné číslo $w \neq 0$. V homogenních souřadnicích je tedy každý bod reprezentován nekonečným počtem vektorů v přidruženém vektorovém prostoru, které jsou ovšem všechny navzájem kolineární. Speciálně můžeme také volit $w = 1$. Dostaneme tak trojici $(x, y, 1)$. Zvláštní význam této volby je patrný z obr. 6.6. Je-li naopak trojice (x, y, w) homogenními souřadnicemi nějakého bodu, pak také trojice $(x/w, y/w, 1)$ je homogenními souřadnicemi téhož bodu a dvojice $(x/w, y/w)$ je jeho souřadnicemi afinními. Projektivní transformace bodu o afinních souřadnicích (x, y) je popsána rovnicí

$$\begin{bmatrix} \varphi(x, y) \\ \psi(x, y) \\ \omega(x, y) \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \quad (6.26)$$



Obr. 6.6: Zavedení homogenních souřadnic

Homogenní souřadnice bodu po transformaci jsou $(\varphi(x, y), \psi(x, y), \omega(x, y))$, odkud pro afinní souřadnice plyne vztah $(\varphi(x, y)/\omega(x, y), \psi(x, y)/\omega(x, y))$. Ze vztahu (6.26) vyplývá, že ve dvojrozměrném prostoru je projektivní transformace je popsána devíti hodnotami $p_{11}, p_{12}, \dots, p_{33}$. Uvedené hodnoty lze ovšem násobit libovolnou nenulovou reálnou konstantou, aniž by se výsledek transformace změnil (násobením se sice změní délka vektoru na levé straně vztahu (6.26), ale bod, který je vektorem reprezentován, je stále tentýž).

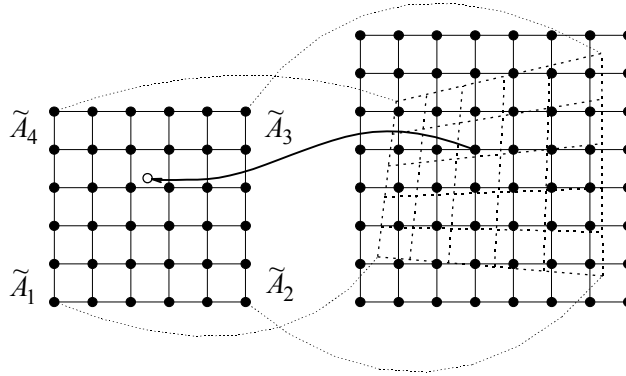
Příklad 6.4. Uvažujme projektivní transformaci zadanou tak, že pro čtveřici bodů $\tilde{A}_1, \tilde{A}_2, \tilde{A}_3, \tilde{A}_4$ ve vstupním obraze jsou předepsány jim odpovídající body A_1, A_2, A_3, A_4 v obraze výstupním (na obr. 6.7 jsou jako body $\tilde{A}_1, \tilde{A}_2, \tilde{A}_3, \tilde{A}_4$ zvoleny rohy obrazu). Naznačíme postup odvození transformační matice ze vztahu (6.26). Souřadnice bodu \tilde{A}_i v původním obraze označíme \tilde{x}_i, \tilde{y}_i . Poloha korespondujícího bodu A_i ve výstupním obraze je předepsána souřadnicemi x_i, y_i . Dále zavedeme označení $\mathbf{p}_1 = (p_{11}, p_{12}, p_{13})^\top$, $\mathbf{p}_2 = (p_{21}, p_{22}, p_{23})^\top$, $\mathbf{p}_3 = (p_{31}, p_{32}, p_{33})^\top$, $\mathbf{x}_i = (x_i, y_i, 1)^\top$. Na základě vztahu (6.26) nyní pro každý ze zadaných bodů dostáváme

$$\begin{bmatrix} \tilde{w}_i \tilde{x}_i \\ \tilde{w}_i \tilde{y}_i \\ \tilde{w}_i \end{bmatrix} = \begin{bmatrix} p_{11}^\top \\ p_{21}^\top \\ p_{31}^\top \end{bmatrix} x_i.$$

Ze třetí rovnice v uvedeném vztahu máme $\tilde{w}_i = \mathbf{p}_3^\top \mathbf{x}_i$. Dosazením do prvních dvou rovnic po úpravě dostaneme

$$\begin{bmatrix} \mathbf{x}_i^\top & \mathbf{0} & -\tilde{x}_i \mathbf{x}_i^\top \\ \mathbf{0} & \mathbf{x}_i^\top & -\tilde{y}_i \mathbf{x}_i^\top \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Vidíme, že každý ze zadaných bodů přispívá k nalezení transformační matice dvěma rovnicemi, což dává celkem osm rovnic pro devět neznámých. Protože víme, že transformační matici je možné násobit libovolným reálným číslem, aniž by se výsledek transformace změnil, zvolíme k zajištění jednoznačnosti řešení doplňující podmínku. Nejsnáze se prakticky realizuje postup, kdy některý z prvků transformační matice položíme roven předem zvolené hodnotě (prosíme čtenáře, aby promyslel úskalí tohoto řešení).



Obr. 6.7: Projektivní transformace obrazu

Nelineární geometrickou transformací obrazu rozumíme transformaci, kde $\varphi(x, y)$, $\psi(x, y)$ jsou nelineární funkce. Nečastěji se jedná o polynomy. Můžeme je volit např. ve tvaru

$$\varphi(x, y) = x + \Delta x(x, y), \quad \psi(x, y) = y + \Delta y(x, y), \quad (6.27)$$

kde

$$\Delta x(x, y) = \sum_{r=0}^m \sum_{s=0}^{m-r} a_{rs} x^r y^s, \quad \Delta y(x, y) = \sum_{r=0}^m \sum_{s=0}^{m-r} b_{rs} x^r y^s \quad (6.28)$$

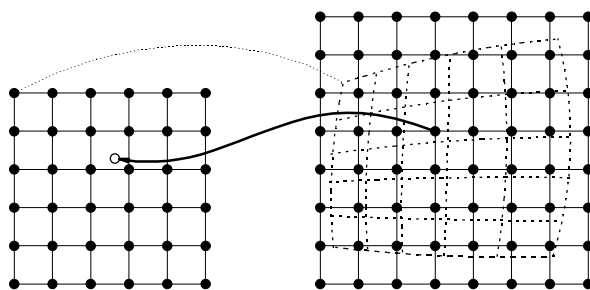
nebo také

$$\Delta x(x, y) = \sum_{r=0}^m \sum_{s=0}^m a_{rs} x^r y^s, \quad \Delta y(x, y) = \sum_{r=0}^m \sum_{s=0}^m b_{rs} x^r y^s. \quad (6.29)$$

Ze vztahu (6.27) je zřejmé, že funkce $\Delta x(x, y)$, $\Delta y(x, y)$ popisují „přemístění“ bodu, ke kterému během transformace dojde. Zavedení funkcí $\Delta x(x, y)$, $\Delta y(x, y)$ je názorné a také výhodné pro praktickou implementaci. Polynomy, které se pro transformaci používají, v praxi zpravidla nebývají příliš vysokého stupně (obvykle druhého nebo třetího). Neznámé hodnoty koeficientů a_{rs} , b_{rs} je možné stanovit tak, že pro jistý počet bodů (x_i, y_i) předepíšeme hodnoty $\varphi(x_i, y_i)$, $\psi(x_i, y_i)$. To dává dvě soustavy rovnic

$$\varphi(x_i, y_i) = x_i + \sum_{r=0}^m \sum_{s=0}^{m-r} a_{rs} x_i^r y_i^s, \quad \psi(x_i, y_i) = y_i + \sum_{r=0}^m \sum_{s=0}^{m-r} b_{rs} x_i^r y_i^s, \quad (6.30)$$

z nichž lze hodnoty a_{rs} , b_{rs} stanovit (v případě polynomu (6.29) lze postupovat analogicky). Předpisy (6.28), (6.29) obsahují $(m+2)(m+1)$ resp. $2(m+1)^2$ neznámých koeficientů. Každá dvojice $[(x_i, y_i), (\varphi(x_i, y_i), \psi(x_i, y_i))]$ dává vzniknout dvěma rovnicím. Je proto zapotřebí nejméně $(m+2)(m+1)/2$ resp. $(m+1)^2$ takových dvojic. Je-li počet dvojic právě roven uvedené hodnotě, obdržíme soustavu lineárních rovnic, kterou lze řešit eliminací. Je-li počet zadaných dvojic větší, obdržíme soustavu, která je předeterminovaná (a dost možná také nekonzistentní).

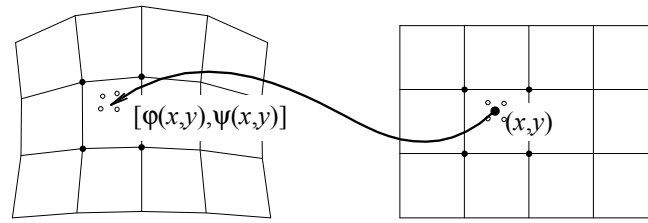


Obr. 6.8: Nelineární geometrická transformace

Je zřejmé, že čím je požadovaná geometrická transformace komplikovanější, tím vyšší stupeň polynomů je nutné v rovnicích (6.28), (6.29) zvolit. Jestliže by stupeň polynomů přesáhl rozumnou mez (řekněme např. 3 až 5), pak je lepší použít postupu, kdy vstupní i výstupní obraz rozdělíme na oblasti a transformační polynomy stanovíme zvlášť pro každou dílčí oblast. Stupeň polynomů může být v tomto případě nízký, protože s ohledem na složitost transformace můžeme volit jemnost dělení. Nejčastěji se proto v tomto případě používají polynomy bilineární. Transformaci využívající dělení na podoblasti lze zadat tak, že výstupní obraz pokryjeme čtyřúhelníkovou sítí a ve vstupním obraze definujeme odpovídající síť transformovanou (obr. 6.9). Síť zadáváme polohou uzlových kontrolních bodů ve vstupním i výstupním obraze. Nejsnáze se implementuje případ, kdy je síť ve výstupním obraze obdélníková. Při praktické realizaci transformace se opět probírají všechny body výstupního obrazu. Podle toho, do které oblasti výstupního obrazu bod padl, se použije odpovídající transformační vztah. Pro i -tou oblast má vztah tvar

$$\varphi_i(x, y) = x + (a_i x + b_i y + c_i xy + d_i), \quad (6.31)$$

$$\psi_i(x, y) = y + (e_i x + f_i y + g_i xy + h_i). \quad (6.32)$$



Obr. 6.9: Geometrická transformace s rozdělením obrazu na oblasti

Neznámé hodnoty $a_i, b_i, c_i, d_i, e_i, f_i, g_i, h_i$ určíme pro každou oblast z podmínky, aby se vrcholy oblasti (uzlové kontrolní body) transformovaly tak, jak bylo předepsáno. Pro každou oblast máme čtyři vrcholy po dvou souřadnicích, což dává celkem osm rovnic potřebných pro určení hledaných hodnot.

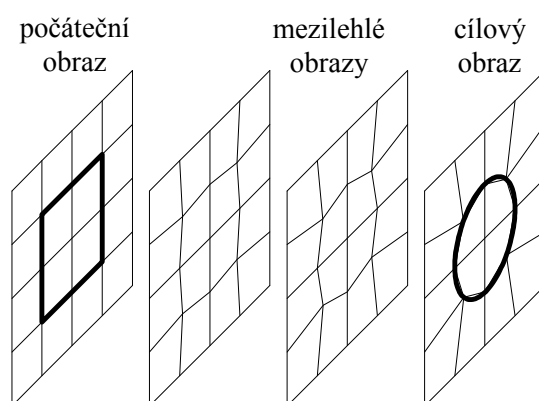
Pro nelineární geometrické operace bývá někdy používáno termínu warping (borečení, kroucení). V televizní tvorbě se v poslední době stala populární také technika nazývaná morphing, která provádí plynulou změnu jednoho objektu v jiný. Předpokládejme, že je zadán počáteční a cílový obraz (označme je $f_s(x, y), f_d(x, y)$). Úkolem je vytvořit sekvenci $f_i(x, y)$ obrazů zachycující postupnou změnu obrazu počátečního v cílový. Jistou možností řešení zadaného úkolu by bylo použít prolínání obrazů podle jednoduchého předpisu

$$f_i(x, y) = (1 - \lambda) f_s(x, y) + \lambda f_d(x, y) \quad (6.33)$$

při hodnotě λ postupně se zvyšující od 0 do 1. Výsledky samotného prolínání však zpravidla nejsou přesvědčivé. Lepšího dojmu lze dosáhnout kombinací prolínání s nelineární geometrickou transformací. K definování transformace lze použít síť kontrolních bodů, kterou pokryjeme počáteční i cílový obraz. Polohu kontrolních bodů předepíšeme tak, aby bylo dosaženo požadované deformace obrazu (obr. 6.10). Síť je nutné pokrýt i obrazy mezilehlé, v nichž lze však síť generovat automaticky pomocí interpolace. Ke generování i -tého mezilehlého obrazu použijeme vztahu

$$f_i(x, y) = (1 - \lambda) f_s[\varphi_{i,s}(x, y), \psi_{i,s}(x, y)] + \lambda f_d[\varphi_{i,d}(x, y), \psi_{i,d}(x, y)], \quad (6.34)$$

kde $\varphi_{i,s}, \psi_{i,s}, \varphi_{i,d}, \psi_{i,d}$ jsou po částech bilineární funkce transformující souřadnice z i -tého mezilehlého obrazu do obrazu počátečního resp. cílového. Uvedené funkce jsou určeny zadanou sítí.



Obr. 6.10: Proměna čtverce na kružnici pomocí morphingu

Kapitola 7

Segmentace obrazu

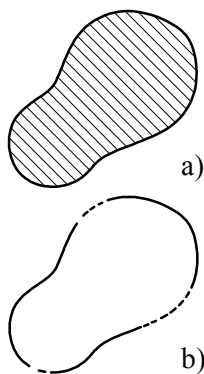
Předpokládejme, že máme obraz, který je popsán obrazovou funkcí, a že máme rozpoznat objekty v tomto obraze. Prvním krokem k tomuto cíli je extrakce objektů, které se v obraze nacházejí. Proces extrakce, v němž jsou objekty separovány navzájem a od nezajímavého pozadí, nazýváme segmentací obrazu. Segmentace je nejčastěji založena na detekci kontur (hran) ohraničujících jednotlivé objekty nebo na detekci celých oblastí, kterými jsou jednotlivé objekty v obraze reprezentovány. Oba uvedené postupy jsou v praxi široce používány, a proto jsou jejich popisu věnovány poměrně rozsáhlé podkapitoly 7.1, 7.3. Podkapitola 7.4 je věnována pohledu, kdy se na segmentaci obrazu pohlíží jako na úlohu o minimalizaci funkcionálu. V podkapitole 7.5 je pak stručně nastíněna problematika zpracování binárních obrazů. To proto, že takové obrazy v procesu segmentace často vznikají.

7.1 Detekce hran

Představujeme si, že každý objekt je v obraze reprezentován souvislou oblastí (obr. 7.1a). Každá oblast je obklopena hranicí. *Hranice* se skládá z hran (případně též z jediné zakřivené hrany). *Hrana* se skládá z jednotlivých hranových bodů. Významným problémem v analýze obrazů je nalezení hran a celých hranic. Při řešení se často ale postupuje tak, že se nejprve naleznou jednotlivé body hran. Uvažujme pro jednoduchost obrazy ve stupních šedi. (I v praxi se před analýzou, a tedy i před hledáním hran, barevné obrazy velmi často převádí na obrazy ve stupních šedi.) Na existenci bodu hrany můžeme usuzovat z průběhu obrazové funkce. Na obr. 7.2 nahoře je znázorněn typický průběh jasu napříč hranou. Za bod hrany se nejčastěji považuje místo, kde průběh jasu vykazuje náhlou změnu, případně inflexní bod. Po nalezení jsou jednotlivé nalezené body hran spojovány různými technikami do hran a celých hranic.

Protože každá plocha by měla být obklopena svojí hranicí (obr. 7.1a), zdá se, že detekce hran a detekce celých ploch by teoreticky měly vést ke vzájemně ekvivalentním výsledkům. V praxi se však mohou vyskytnout problémy. Jsou-li např. rozpoznávané objekty v obraze špatně zřetelné nebo je-li obraz poškozen šumem, pak

nemusí nalezené hrany tvořit kolem objektů uzavřené smyčky, jak si teoreticky představujeme (obr. 7.1b). Vlivem šumu mohou být na druhé straně hrany detekovány i tam, kde ve skutečnosti nejsou.



Obr. 7.1: Detekce oblastí stanovením hranice

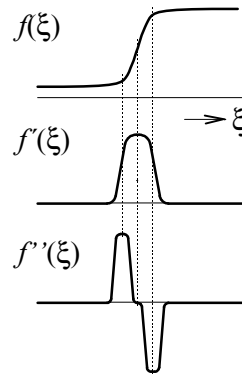
7.1.1 Gradientní metody hledání hran

Obr. 7.2 ukazuje typický průběh jasu ve směru napříč hranou a první a druhou derivaci tohoto průběhu. Gradientní metody využívají skutečnosti, že v místě hrany má absolutní hodnota první derivace průběhu jasu vysokou hodnotu. Hodnota derivace popisuje intenzitu kontury v daném bodě - hovoříme o velikosti hrany. Operátory, které velikost hrany stanovují (obvykle se aplikují v každém bodě obrazu), bývají nazývány hranovými operátory. Z uvedeného vyplývá, že nejjednoduššími hranovými operátory jsou zřejmě derivace $\partial f/\partial x$ a $\partial f/\partial y$, které popisují změnu úrovně jasu ve směru os x a y . Těchto operátorů by bylo možné použít k hledání hran rovnoběžných se souřadnými osami. Při hledání hran obecného směru je zapotřebí vyšetřovat průběh jasu ve směru kolmém na směr potenciální hrany. Nechť vektor $\mathbf{n} = (\cos \theta, \sin \theta)$ popisuje směr kolmý ke směru eventuální hrany a ξ nechť je souřadnice měřená v tomto směru. K ověření skutečnosti, zda v daném bodě existuje hrana známého směru, lze použít derivaci ve směru. Pro derivaci ve směru kolmo k hraně máme

$$\frac{\partial f}{\partial \xi} = \text{grad}(f) \cdot \mathbf{n} = \frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta. \quad (7.1)$$

Za velikost hrany v daném bodě můžeme považovat hodnotu $|\partial f/\partial \xi|$.

Při hledání hrany její směr předem obvykle neznáme. Rovnice (7.1) a úvahy, které k ní vedly, však podávají návod, jak situaci řešit. Pro rozhodnutí, zda ve vyšetřovaném bodě je či není hrana, bude rozhodující směr, v němž je změna jasu největší. Tímto směrem je gradient obrazové funkce. Směr hrany je pak kolmý ke směru gradientu. Za velikost hrany lze vzít velikost gradientu. Označme $e(x, y)$ velikost hrany,

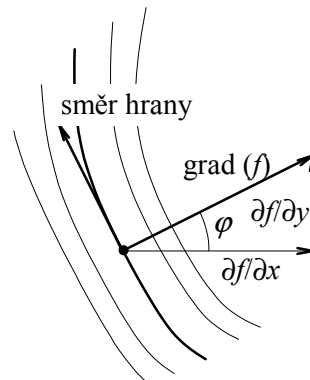


Obr. 7.2: Průběh jasu a jeho první a druhé derivace ve směru napříč hranou

$\varphi(x, y)$ směr gradientu a $\psi(x, y)$ směr hrany v bodě (x, y) . Dále pro stručnost zavedme označení $f_x(x, y) = \partial f(x, y)/\partial x$, $f_y(x, y) = \partial f(x, y)/\partial y$. S využitím právě zavedeného pak máme

$$e(x, y) = \sqrt{f_x^2(x, y) + f_y^2(x, y)}, \quad (7.2)$$

$$\varphi(x, y) = \arctan \left[\frac{f_y(x, y)}{f_x(x, y)} \right], \quad \psi(x, y) = \varphi(x, y) + \frac{\pi}{2}. \quad (7.3)$$



Obr. 7.3: Stanovení velikosti a směru hrany na základě gradientu

Cílem praktického výpočtu zpravidla bývá rozhodnout, zda vyšetřovaný bod obrazu je či není bodem ležícím na hranici nějakého objektu. V nejjednodušším případě lze za body hranice považovat ta místa, kde hodnota $e(x, y)$ je větší než nějaká pevná, předem zvolená prahová hodnota. Tento jednoduchý postup má ovšem své nedostatky: 1) Hranice objektu obvykle vyjde širší než jeden bod (intuitivně si možná představujeme, že tloušťka hranice je právě jeden bod). 2) Uvedený postup neřeší

problém doplnění chybějících a odstranění nadbytečných částí hranice diskutovaný v úvodu podkapitoly 7.1. Důmyslnější postupy popíšeme postupně později (např. v souvislosti s Cannyho detektorem hran).

Až dosud jsme předpokládali, že obrazová funkce je spojitá. Při praktické implementaci pracujeme ovšem nejčastěji s funkcí diskrétní. Jestliže nahradíme derivace diferencemi, můžeme vztahů (7.2), (7.3) snadno použít i pro diskrétní případ. Pro difference můžeme psát

$$f_x(x, y) = f(x + 1, y) - f(x, y), \quad (7.4)$$

$$f_y(x, y) = f(x, y + 1) - f(x, y). \quad (7.5)$$

Při praktické realizaci výpočtu je hodnota $e(x, y)$ stanovována a porovnávána s hodnotou prahu ve všech bodech (pixelech) obrazu. Poznamenejme, že při stanovení délky gradientu lze použít i jiných metrik, než je metrika uvedená ve vztahu (7.2) (kořeny těchto snah sahají ovšem poněkud do minulosti, kdy byl výpočet odmocniny časově velmi náročnou, a proto obávanou operací). Můžeme použít např. následujících vztahů:

$$e(x, y) = |f_x(x, y)| + |f_y(x, y)|, \quad e(x, y) = \max \{|f_x(x, y)|, |f_y(x, y)|\}. \quad (7.6)$$

Skutečnost, že při praktickém výpočtu na použité metrice příliš nezáleží, ukazuje následující příklad, který ponecháváme čtenáři k samostatnému řešení: Dokažte, že pro každou dvojici nezáporných čísel a, b platí $(a + b)/\sqrt{2} \leq \sqrt{a^2 + b^2} \leq \sqrt{2} \max\{a, b\}$.

V minulosti bylo realizováno mnoho praktických metod pro detekci hran, jejichž společným teoretickým základem je výpočet velikosti gradientu. Některé z nich popíšeme dále podrobněji.

Robertsův operátor: Tento hranový operátor je jedním z nejstarších, ale stále ještě občas používaným operátorem pro stanovení velikosti hrany. Operátor stanovuje difference ve dvou na sebe kolmých diagonálních směrech. Velikost hrany v bodě (x, y) obrazu se počítá podle předpisu

$$e(x, y) = \sqrt{(f(x, y) - f(x + 1, y + 1))^2 + (f(x + 1, y) - f(x, y + 1))^2}. \quad (7.7)$$

Operátor Prewittové: Tento operátor využívá k výpočtu velikosti hrany v bodě o souřadnicích x, y hodnot obrazové funkce ve všech sousedních pixelech. Tyto hodnoty pro stručnost označíme A, B, C, D, F, G, H, I (obr. 7.4). Operátor počítá derivace ve směrech os x, y průměrem

$$f_x(x, y) = \frac{1}{3} [(C - A) + (F - D) + (I - G)],$$

$$f_y(x, y) = \frac{1}{3} [(A - G) + (B - H) + (C - I)]. \quad (7.8)$$

Uvedený postup stanovení derivací průměrem přispívá k redukci šumu. V případech, kdy nezáleží na absolutní velikosti hrany, lze koeficienty $1/3$ vypustit. Jestliže po výpočtu velikosti hrany následuje prahování, v němž se rozhoduje, zda v daném bodě hrana existuje či nikoli, lze vypuštění koeficientů kompenzovat volbou velikosti prahu.

A	B	C
D	$f(x, y)$	F
G	H	I

Obr. 7.4: Značení hodnot obrazové funkce pro operátor Prewittové

Sobelův operátor: Tento operátor je podobný operátoru Prewittové. Rozdíl spočívá v tom, že se pro výpočet derivací v jednotlivých směrech používá vážený průměr dle vztahů

$$f_x(x, y) = \frac{1}{4} [(C - A) + 2(F - D) + (I - G)] , \quad (7.9)$$

$$f_y(x, y) = \frac{1}{4} [(A - G) + 2(B - H) + (C - I)] . \quad (7.10)$$

I v tomto případě lze za stejných okolností jako u operátoru Prewittové koeficienty $1/4$ vypustit.

Kirschův operátor: Také tento operátor lze zařadit mezi operátory založené na výpočtu gradientu. Veličina mající povahu derivace je počítána celkem v osmi směrech a za velikost hrany je vzata její maximální hodnota. Pro stručnost přeznačíme funkční hodnoty obrazové funkce tak, jak je uvedeno na obr. 7.5. Velikost hrany v bodě (x, y) se vypočítá dle vzorce

$$e(x, y) = \max_{i=0}^7 (|5S_i - 3T_i|) , \quad (7.11)$$

kde

$$S_i = A_i + A_{i+1} + A_{i+2} , \quad (7.12)$$

$$T_i = A_{i+3} + A_{i+4} + A_{i+5} + A_{i+6} + A_{i+7} . \quad (7.13)$$

Index i , pro který je ve výrazu (7.11) dosaženo maxima, určuje směr hrany. Jestliže ve výrazech (7.12), (7.13) vyjdou hodnoty indexů větší než 7, upraví se operací modulo 8.

7.1.2 Detekce hran hledáním průchodu druhé derivace nulou

Z obr. 7.2 je zřejmé, že druhá derivace průběhu jasu ve směru napříč hranou nabývá dvou extrémů opačného znaménka a v místě hrany znaménko mění. Protože směr

A_0	A_1	A_2
A_7	$f(x, y)$	A_3
A_6	A_5	A_4

Obr. 7.5: Značení hodnot obrazové funkce pro Kirschův operátor

hrany není zpravidla předem znám, je nutné druhou derivaci počítat přinejmenším ve dvou směrech, např. ve směru souřadných os x , y . K výpočtu lze použít Laplaceova operátoru. Zavedme označení $f_{xx}(x, y) = \partial^2 f(x, y)/\partial x^2$, $f_{yy}(x, y) = \partial^2 f(x, y)/\partial y^2$. Laplaceův operátor je definován vztahem

$$\nabla^2 f(x, y) = f_{xx}(x, y) + f_{yy}(x, y). \quad (7.14)$$

Je-li obrazová funkce diskrétní, použijeme místo derivací diference. Pro body uvnitř obrazu (nikoli v krajních řádcích a sloupcích) můžeme použít symetrických diferenčních vzorců (pro čtenáře by mělo být snadné si je odvodit)

$$f_{xx}(x, y) = f(x - 1, y) - 2f(x, y) + f(x + 1, y), \quad (7.15)$$

$$f_{yy}(x, y) = f(x, y - 1) - 2f(x, y) + f(x, y + 1). \quad (7.16)$$

Dosazením vztahů (7.15), (7.16) do (7.14) dostaneme pro realizaci Laplaceova operátoru předpis

$$\nabla^2 f(x, y) = f(x + 1, y) + f(x - 1, y) + f(x, y + 1) + f(x, y - 1) - 4f(x, y). \quad (7.17)$$

Vztah (7.17) lze vypočítat pomocí konvoluce obrazové funkce $f(x, y)$ s konvoluční maskou uvedenou na obr. 7.6. V souvislosti s určováním hran pomocí druhé derivace bývá uváděna také maska dle obr. 8.6b, která realizuje výpočet součtu druhých derivací ve čtyřech směrech 0° , 45° , 90° , 135° . Je zřejmé, že v krajních řádcích a sloupcích obrazu nelze masku z obr. 7.6 použít. Pro uvedené případy lze v případě potřeby snadno odvodit speciální masky vyžadující pouze hodnoty uvnitř obrazu. Např. pro výpočet hodnoty Laplaceova operátoru v dolním řádku obrazu snadno odvodíme masku dle obr. 7.7a, pro výpočet hodnoty v levém dolním rohu pak masku uvedenou na obr. 7.7b (ostatní potřebné masky získáme rotací masek zde uvedených). Hrany v obraze nalezneme analýzou výsledku poskytnutého Laplaceovým operátorem. Za hranu považujeme místo, kde mezi dvěma dostatečně velikými a dostatečně blízko položenými extrémy opačného znaménka výsledek poskytnutý operátorem mění znaménko.

Právě popsaná metoda hledání hran založená na výpočtu druhé derivace a na použití Laplaceova operátoru je dosti citlivá na šum. I při malém zašumění obrazu je detekováno značné množství falešných hran. Problém lze řešit filtrací šumu, která se provede ještě před tím, než jsou hledány hrany. Marr a Hildreth [12] popsali metodu,

0	1	0	0	1	0
1	-3	1	0	-2	1
0	0	0	0	0	0
a)			b)		

Obr. 7.6: Konvoluční masky pro výpočet Laplaceova operátoru

0	1	0	1	1	1
1	-4	1	1	-8	1
0	1	0	1	1	1
a)			b)		

Obr. 7.7: Příklady masek pro body na krajích a v rozích obrazu

v níž pro redukci šumu používají konvoluci s gausiánem. Protože se tato metoda používá dosti často, popíšeme ji podrobněji. Dvourozměrný gausián je definován vztahem (obr. 7.7)

$$G(x, y) = G(x)G(y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right). \quad (7.18)$$

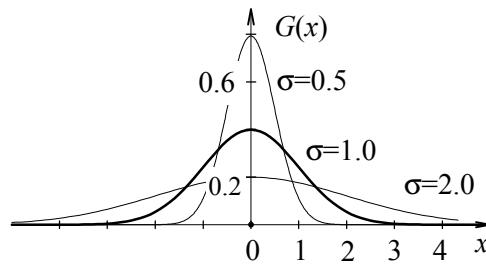
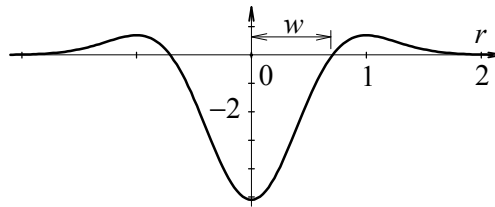
Hodnotou σ lze „regulovat šířku“ gausiánu a tudíž i míru filtrace. Běžně používané hodnoty σ jsou např. 0.5 až 3. Realizaci konvoluce obrazové funkce s gausiánem a následnou aplikaci Laplaceova operátoru popisují vztahy

$$\begin{aligned} & \nabla^2 [G(x, y) * f(x, y)] & (7.19) \\ & = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \int \int_{-\infty}^{\infty} G(x - \xi, y - \eta) f(\xi, \eta) d\xi d\eta \\ & = \left[\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) G(x, y) \right] * f(x, y) \\ & = [\nabla^2 G(x, y)] * f(x, y). \end{aligned}$$

Poslední z uvedených vztahů ukazuje, že popisovanou posloupnost akcí lze realizovat konvolucí obrazové funkce s funkcí $\nabla^2 G$. Položíme-li $r = \sqrt{(x^2 + y^2)}$, můžeme funkci $\nabla^2 G$ zapsat ve tvaru

$$\nabla^2 G(r) = \left(\frac{r^2 - 2\sigma^2}{2\pi\sigma^6} \right) \exp\left(\frac{-r^2}{2\sigma^2}\right). \quad (7.20)$$

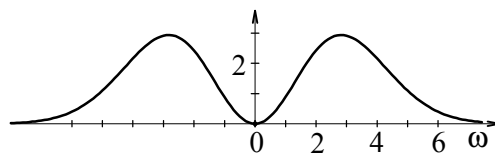
Funkce $\nabla^2 G$ je rotačně symetrická. Její průběh (průřez průběhu) je znázorněn na obr. 7.9. Označme w poloměr kružnice, kde funkce přechází ze záporných do kladných hodnot (obr. 7.9). Z rovnice (7.20) vychází $w = \sigma\sqrt{2}$.

Obr. 7.8: Průběh gausiánu $G(x)$ pro $\sigma = 0.5$, $\sigma = 1$, $\sigma = 2$ Obr. 7.9: Průběh funkce $\nabla^2 G(r)$, $\sigma = 0.5$

K důkladnějšímu osvětlení principu činnosti této metody vyšetříme ještě frekvenční charakteristiku operátoru $\nabla^2 G$. Provedením Fourierovy transformace a položením $u^2 + v^2 = \omega^2$ dostaneme

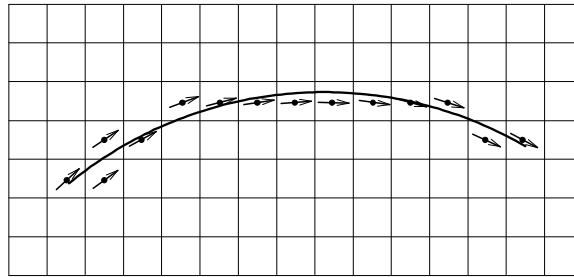
$$\mathcal{F} \{ \nabla^2 G(x, y) \} = \omega^2 \exp \left(-\sigma^2 \frac{\omega^2}{2} \right). \quad (7.21)$$

Průběh frekvenční charakteristiky $\mathcal{F} \{ \nabla^2 G \}$ je znázorněn na obr. 7.10. Je zřejmé, že se jedná o pásmový filtr. Vrchol frekvenční charakteristiky vychází v místě $\omega = \sqrt{2}/\sigma$. Po provedení konvoluce $[\nabla^2 G(x, y)] * f(x, y)$ pokračuje metoda hledáním průchodů nulou stejně, jak již bylo popsáno dříve.

Obr. 7.10: Frekvenční charakteristika $\mathcal{F} \{ \nabla^2 G \}$ pro $\sigma = 0.5$

7.2 Spojování hran

Hranové operátory z předchozí podkapitoly určovaly velikost a případně i směr hrany jednotlivě v každém bodě obrazu, tedy lokálně. To je ovšem poněkud v rozporu s intuitivní představou hran jako delších souvislých úseček případně křivek v obrazech. Hrany chápané tímto druhým způsobem nazýváme v této podkapitole globálními hranami. K nalezení globálních hran lze využít výsledků poskytovaných lokálními operátory, které jsou však dále zpracovány v procesu tzv. spojování hran. Proces spojování je ilustrován na obr. 7.11. Ve vyobrazeném případě umožňují lokálně stanovené velikosti a směry hran snadno proložit hranu globální. Lze k tomu použít obvyklých technik jako je např. prokládání úseček, oblouků či jiných útvarů metodou nejmenších čtverců. Při praktické realizaci se ale zpravidla bohužel nevyhneme různým heuristickým postupům. Je to hlavně proto, že prokládaných útvarů je obvykle více a je zapotřebí rozhodnout, které mají být proloženy kterými hranovými body, což je obtížná úloha. My se v této podkapitole omezíme pouze na jednu elegantní a velmi často citovanou i používanou metodu, kterou je Houghova transformace.



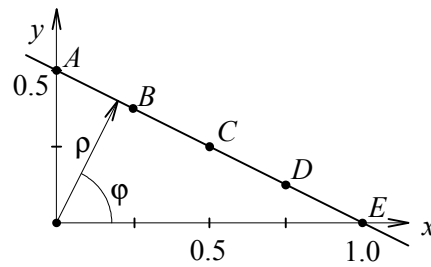
Obr. 7.11: Proložení globální hrany

7.2.1 Houghova transformace

Předpokládejme, že jsme již pomocí některé z metod z podkapitoly 7.1 rozhodli, které body obrazu považujeme za body hran. Nyní chceme ověřit, zda jsou body hran v obraze položeny tak, že naznačují existenci přímek (úseček) v obraze. Metoda, kterou zde popíšeme, byla navržena Houghem a později zdokonalena Dudou a Hartem [8]. Metoda je známa pod názvem *Houghova transformace*. Uvažujme rovnici přímky ve tvaru (obr. 7.12)

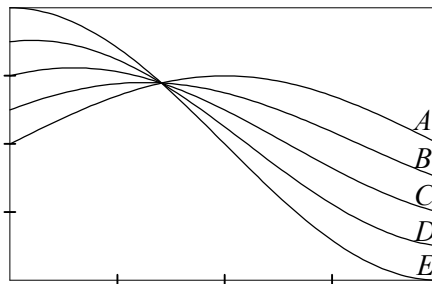
$$\rho = x \cos \varphi + y \sin \varphi. \quad (7.22)$$

Pro přímku procházející bodem Q reprezentovaným vektorem (x_Q, y_Q) pak platí $\rho = x_Q \cos \varphi + y_Q \sin \varphi$. Uvažujme nyní prostor φ, ρ . V tomto prostoru odpovídá každé přímce procházející bodem Q jediný bod. Svazku přímek procházejících bodem Q pak v tomto prostoru odpovídá sinusoida $\rho = x_Q \cos \varphi + y_Q \sin \varphi$. Mějme nyní



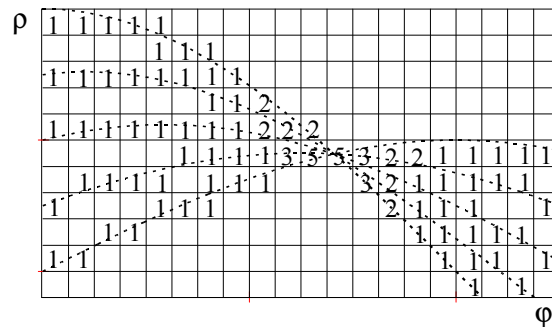
Obr. 7.12: Houghova transformace: Body ležící na přímce

v obraze např. body A, B, C, D, E , které leží na přímce (obr. 7.12). Svazkům možných přímek v těchto bodech odpovídají v prostoru φ, ρ sinusoidy A, B, C, D, E (obr. 7.13). Obrazem přímky, na které leží body A, B, C, D, E (tedy přímky, která je společná všem svazkům), je v prostoru φ, ρ bod, v němž se sinusoidy A, B, C, D, E protínají.

Obr. 7.13: Bodům ležícím na přímce odpovídá v prostoru φ, ρ průsečík sinusoid

Dosud provedené úvahy vedou k následujícímu postupu výpočtu. Rozdělme prostor φ, ρ na oblasti. Nechť např. dvojice i, j označuje v prostoru φ, ρ obdélníkovou oblast, pro niž platí $\varphi_{i-1} \leq \varphi < \varphi_i, \rho_{j-1} \leq \rho < \rho_j$. Zavedme nad takto diskretizovaným prostorem φ, ρ dvourozměrný histogram $h(i, j)$. Počáteční hodnota $h(i, j)$ je pro všechna (i, j) nula. Nyní probírejte všechny pixely obrazu. Pro každý pixel $Q = (x_Q, y_Q)$, kde v obraze nalezneme hranu (předpokládáme, že neznáme její směr), sestrojíme v prostoru φ, ρ sinusoidu (obr. 7.14). Při tom v histogramu inkrementujeme hodnoty $h(i, j)$ pro ty oblasti (i, j) , kterými sinusoida prochází. Po skončení tohoto procesu $h(i, j) = n$ popisuje skutečnost, že na přímce charakterizované hodnotami $\varphi \in \langle \varphi_{i-1}, \varphi_i \rangle, \rho \in \langle \rho_{j-1}, \rho_j \rangle$ leží v obraze n bodů, které byly považovány za hrany. Pomocí Houghovy transformace tak můžeme zjistit, že v obraze leží v jedné přímce vysoký počet bodů hran. V případě potřeby je toto zjištění možné dále doplnit analýzou, jak jsou detekované body rozloženy podél nalezené přímky.

Jestliže v jednotlivých bodech obrazu dokážeme hranu nejen detekovat, ale také



Obr. 7.14: Bodům ležícím na přímce odpovídá vysoká hodnota v histogramu

dostatečně spolehlivě určit její směr, pak lze popsanou metodu zjednodušit. Necht x, y je bod hrany a ψ směr hrany v tomto bodě. Potom podle vztahu $\rho = x \cos(\psi - \pi/2) + y \sin(\psi - \pi/2)$ můžeme ihned vypočítat odpovídající hodnotu ρ . V tomto případě tedy místo generování sinusoidy generujeme pouze jednotlivé body a histogram inkrementujeme pouze v jediné oblasti. Na závěr poznamenejme, že Houghovu transformaci lze zobecnit také na křivky. Protože počet hledaných parametrů křivky udává potřebný rozměr histogramu, jedná nejčastěji o křivky se dvěma, nejvýše třemi parametry.

7.2.2 Cannyho detektor hran

Popisovaný detektor hran byl publikován Cannym [4], [5] a později detailně studován mnoha dalšími autory. Při návrhu detektoru vycházel Canny z toho, že starší detektory hran byly konstruovány více či méně intuitivně, a formuloval proto požadavky, které by měl detektor splňovat. K nalezení detektoru pak přistoupil jako k optimalizační úloze (tj. hledal detektor, který vytyčené požadavky splňuje co nejlépe). Cannym navržené požadavky byly následující: 1) minimalizovat pravděpodobnost chybné detekce, 2) najít polohu hrany v obraze co nejpřesněji, 3) bod hrany identifikovat jednoznačně. Postup, který vede k návrhu odpovídajícího filtru, popíšeme podrobněji. Budeme možná překvapeni, že naznačený důmyslný přístup povede nakonec k detektoru gradientního typu. Poznamenejme ještě, že odvození Cannyho detektoru je možná poměrně zdoluhavé. Je však zajímavé a poučné. Cannyho detektor bývá dnes již standardní součástí každé knihovny pro zpracování obrazu. K využití detektoru není zapotřebí znát postup odvození podrobně. Znalost výchozích principů ale umožní používat detektoru s větší jistotou.

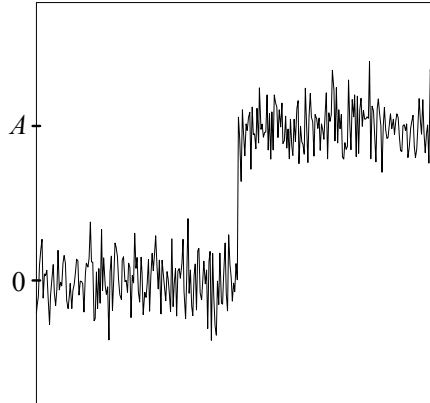
Předpokládejme nejprve, že vstupní signál je jednorozměrný (získané výsledky později zobecníme i na signály dvojrozměrné). Matematickým modelem hrany ve vstupním signálu (ve shodě s původním pramenem označíme vstupní signál $I(x)$) je skok výšky A umístěný do počátku. Doplněn je dále aditivní šum $n(x)$ (obr. 7.15). Předpokládáme, že šum je bílý s nulovou střední hodnotou. Vstupní signál má tedy

tvar

$$I(x) = Au_{-1}(x) + n(x), \quad \text{kde} \quad u_{-1}(x) = \begin{cases} 0 & \text{pro } x < 0 \\ 1 & \text{pro } x \geq 0 \end{cases}. \quad (7.23)$$

Předpokládáme, že hranový operátor bude založen na výpočtu konvoluce vstupního signálu s funkcí $f(x)$. Tuto funkci, kterou zatím neznáme, stanovíme tak, aby byla co nejlépe splněna kriteria formulovaná v úvodu této podkapitoly. Označme $O(x_0)$ signál, který je výsledkem konvoluce. Podle definice konvoluce máme

$$O(x_0) = \int_{-\infty}^{\infty} I(x_0 - x)f(x) dx. \quad (7.24)$$



Obr. 7.15: Průběh signálu ve směru napříč hranou

Nejprve vyhodnotíme, jaký výstup v bodě $x_0 = 0$ odpovídá skoku velikosti A (bez šumu). Dosazením rovnice (7.23) do rovnice (7.24) a položením $n(x) = 0, x_0 = 0$ dostaneme

$$\int_{-\infty}^{\infty} Au_{-1}(-x)f(x) dx = A \int_{-\infty}^0 f(x) dx. \quad (7.25)$$

Dále vypočítáme střední hodnotu čtverce signálu, který na svém výstupu poskytuje operátor v bodě $x_0 = 0$, jestliže je na vstup přiložen pouze šum $n(x)$. Dostaneme

$$E \left\{ \left[\int_{-\infty}^{\infty} n(-x)f(x) dx \right]^2 \right\} = \int_{-\infty}^{\infty} E \{ n^2(-x) \} f^2(x) dx = n_0^2 \int_{-\infty}^{\infty} f^2(x) dx. \quad (7.26)$$

Při úpravě výše uvedeného vztahu jsme se opírali o předpoklad, že $n(x)$ je bílý šum s nulovou střední hodnotou a že tedy pro všechna $x_1 \neq x_2$ platí $E\{n(x_1)n(x_2)\} = 0$.

Dále jsme předpokládali, že hodnota $E\{n^2(x)\}$ je pro všechna x konstantní, a položili jsme proto $E\{n^2(x)\} = n_0^2$ (jedná se o disperzi šumové složky vstupního signálu). Definujme poměr signál/šum (SNR) jako podíl odezvy na skok velikosti A ku odmocnině střední hodnoty čtverce odezvy na šum. Pro hodnocení pravděpodobnosti chybné detekce dále zavedme míru $\Sigma = n_0 SNR/A$, která je závislá pouze na hledané funkci $f(x)$, nikoli na velikosti vstupního signálu. Máme tedy

$$SNR = \frac{A \int_{-\infty}^0 f(x) dx}{n_0 \sqrt{\int_{-\infty}^{\infty} f^2(x) dx}}, \quad \Sigma = \frac{n_0}{A} SNR = \frac{\int_{-\infty}^0 f(x) dx}{\sqrt{\int_{-\infty}^{\infty} f^2(x) dx}}. \quad (7.27)$$

Při detekci hrany právě konstruovaným hranovým operátorem budeme za hranu považovat místo, kde operátor (konvoluce s f) dává největší hodnotu. V případě hrany umístěné do počátku by to teoreticky mělo být místo $x_0 = 0$. Šum však způsobí, že tomu tak nebude zcela přesně. Nalezneme proto místo, kde výstup operátoru nabývá extrému. K tomu vypočítáme první derivaci výstupního signálu a položíme ji rovnu nule

$$\frac{d}{dx_0} O(x_0) = O'(x_0) = \frac{d}{dx_0} \int_{-\infty}^{\infty} I(x_0 - x) f(x) dx = 0. \quad (7.28)$$

Pro derivování konvoluce máme

$$\begin{aligned} \frac{d}{dx_0} \int_{-\infty}^{\infty} I(x_0 - x) f(x) dx &= \frac{d}{dx_0} \int_{-\infty}^{\infty} I(x) f(x_0 - x) dx \\ &= \int_{-\infty}^{\infty} I(x) f'(x_0 - x) dx = \int_{-\infty}^{\infty} I(x_0 - x) f'(x) dx. \end{aligned} \quad (7.29)$$

Výstup rozdělíme na složku odpovídající skoku umístěnému do počátku a na složku odpovídající šumu. Tyto složky označíme postupně O_s , O_n . Derivace odezvy na skok má v bodě x_0 velikost

$$O'_s(x_0) = \int_{-\infty}^{\infty} A u_{-1}(x_0 - x) f'(x) dx = A \int_{-\infty}^{x_0} f'(x) dx = A f(x_0). \quad (7.30)$$

Derivaci $O'_s(x_0)$ odezvy na skok rozvineme kolem počátku v Taylorovu řadu. Předpokládáme při tom, že hledaná funkce $f(x)$ bude antisymetrická, a že proto bude platit $f(0) = 0$ (pokud by to nebyla pravda, řešení bychom nenalezli). Dostaneme

$$O'_s(x_0) = A f(x_0) = A f(0) + x_0 A f'(0) = x_0 A f'(0). \quad (7.31)$$

Derivací odezvy hranového operátoru na šum bude náhodná proměnná s nulovou střední hodnotou a s variancí dle vztahu

$$E \{ O_n'^2(x_0) \} = E \left\{ \left[\int_{-\infty}^{\infty} n(x_0 - x) f'(x) dx \right]^2 \right\} = n_0^2 \int_{-\infty}^{\infty} f'^2(x) dx. \quad (7.32)$$

Pro podmínku extrému pak máme

$$O'(x_0) = O_s'(x_0) + O_n'(x_0) = 0 \quad \text{a tedy} \quad O_s'(x_0) = -O_n'(x_0). \quad (7.33)$$

Z předchozí rovnice pak dále plyne

$$E \{ O_s'^2(x_0) \} = E \{ O_n'^2(x_0) \}. \quad (7.34)$$

Dosazením vztahů (7.31) a (7.32) do rovnice (7.34) dostaneme:

$$E \{ x_0^2 \} = \frac{n_0^2 \int_{-\infty}^{\infty} f'^2(x) dx}{A^2 f'^2(0)} = \sigma_{x_0}^2. \quad (7.35)$$

V předchozím výraze jsme označení σ_{x_0} použili pro aproximaci směrodatné odchylky polohy maxima na výstupu operátoru od skutečné polohy skoku, který leží v počátku. Schopnost operátoru správně lokalizovat hranu hodnotíme podle převrácené hodnoty této odchylky. Zavedeme míru $\Lambda = n_0/(\sigma_{x_0}A)$, která závisí pouze na vlastnostech hranového operátoru a nikoli na signálu. S využitím vztahu (7.35) dostaneme

$$\Lambda = \frac{|f'(0)|}{\sqrt{\int_{-\infty}^{\infty} f'^2(x) dx}}. \quad (7.36)$$

Výkonnost hranového operátoru pak hodnotíme součinem $\Sigma\Lambda$. S využitím vztahů (7.27), (7.36) máme

$$\Sigma\Lambda = \frac{\int_{-\infty}^0 f(x) dx}{\sqrt{\int_{-\infty}^{\infty} f^2(x) dx}} \frac{|f'(0)|}{\sqrt{\int_{-\infty}^{\infty} f'^2(x) dx}}. \quad (7.37)$$

Poslední podmínkou je, aby jedné hraně odpovídalo pokud možno jediné maximum na výstupu operátoru. Tento požadavek vyjádříme jako požadavek na funkci $f(x)$. Využijeme pomocného tvrzení, které uvedeme bez důkazu a které říká, že střední vzdálenost x_{zc} mezi dvěma průchody nulou odezvy funkce g na gaussovský šum je dána výrazem (odezvou se rozumí konvoluce funkce g s šumem)

$$x_{zc} = \pi \left(\frac{-R(0)}{R''(0)} \right)^{\frac{1}{2}}, \quad (7.38)$$

kde $R(\tau)$ je autokorelační funkce funkce g . Snadno ověříme, že platí (předpokládáme $g(-\infty) = g(\infty) = 0$)

$$R(0) = \int_{-\infty}^{\infty} g^2(x) dx, \quad R''(0) = \int_{-\infty}^{\infty} g(x)g''(x) dx = - \int_{-\infty}^{\infty} g'^2(x) dx. \quad (7.39)$$

V našem případě hledáme střední vzdálenost x_{zc} mezi dvěma průchody nulou odezvy funkce f' . Ze vztahů (7.38), (7.39) dostaneme

$$x_{zc} = \pi \left(\frac{\int_{-\infty}^{\infty} f'^2(x) dx}{\int_{-\infty}^{\infty} f''^2(x) dx} \right)^{\frac{1}{2}}. \quad (7.40)$$

Vzdálenost x_{\max} mezi dvěma přilehlými maximy odezvy funkce na šum je dvojnásobkem x_{zc} . Tuto hodnotu vyjádříme relativně vzhledem k šířce operátoru. Šířkou operátoru rozumíme interval $\langle -W, W \rangle$. Na tomto intervalu předpokládáme funkční hodnoty f nenulové (mimo tento interval předpokládáme funkční hodnoty nulové; dále připomeňme, že funkci f předpokládáme antisymetrickou). Je tedy

$$x_{\max} = 2x_{zc} = kW. \quad (7.41)$$

Schopnost detektoru identifikovat hranu jednoznačně bude tím lepší, čím vyšší bude hodnota k . Dosud neznámá funkce f by tedy měla maximalizovat výraz (7.37) a při tom dávat pokud možno co nejvyšší hodnotu k . V [4] je popsáno řešení založené na maximalizaci funkcionálu (7.37). Jsou postupně voleny různé hodnoty k (0.075 až 0.7) a rovnice (7.41) je použita jako doplňující podmínka. Pro každou hodnotu k je numerickou optimalizací stanovena funkce f . Na základě takto získaných výsledků je ukázáno, že pro vyšší hodnoty k (což je prakticky zajímavý případ) lze hledanou funkci f dosti dobře aproximovat pomocí první derivace gausiánu. Na tuto možnost se dále zaměříme. Připomeňme, že gausián je definován předpisem

$$G(x) = \exp\left(-\frac{x^2}{2\sigma^2}\right). \quad (7.42)$$

Funkce f má tedy tvar

$$f(x) \approx G'(x) = -\frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right). \quad (7.43)$$

Pro výše uvedenou aproximaci funkce f vypočteme následující hodnoty [4]

$$|f'(0)| = \frac{1}{\sigma^2}, \quad \int_{-\infty}^0 f(x) dx = 1, \quad \int_{-\infty}^{\infty} f^2(x) dx = \frac{\sqrt{x}}{2\sigma}, \quad (7.44)$$

$$\int_{-\infty}^{\infty} f''(x) dx = \frac{3\sqrt{\pi}}{4\sigma^3}, \quad \int_{-\infty}^{\infty} f''(x) dx = \frac{15\sqrt{\pi}}{8\sigma^5}.$$

Odtud pak index $\Sigma\Lambda$ výkonnosti operátoru a hodnota k pro operátor $f(x) = G'(x)$ vychází

$$\Sigma\Lambda = \sqrt{\frac{8}{3\pi}} = 0.92 \quad \text{a} \quad k = \sqrt{\frac{4}{15}} = 0.51. \quad (7.45)$$

Zjištěné hodnoty jsou jen o málo menší než hodnoty pro operátor optimální. Podstatnou výhodou použití gausiánu je ale snadný výpočet.

Až doposud jsme se soustředili na jednorozměrný případ. Zbývá ukázat, jak pracuje Cannyho detektor v případě dvojrozměrném. Předpokládejme nejprve, že máme hledat hrany známého směru. Ve směru napříč hranou použijeme konvoluci s právě nalezenou funkcí f . Ve směru podél hrany uplatníme projekci. Jestliže je f aproximována derivací gausiánu, pak je výhodné, aby funkcí realizující projekci byl rovněž gausián (projekce bude počítána jako konvoluce s gausiánem). Jako obvykle není zapotřebí výpočet provádět pro všechny možné směry hran, ale postačí jej provést pro dva směry (např. x, y). Nechť $I(x, y)$ je obrazová funkce. Pro hrany rovnoběžné s osou y a hrany rovnoběžné s osou x pak dává hranový operátor následující hodnoty

$$[I(x, y) * G(y)] * G_x(x) = \frac{\partial}{\partial x} \{I(x, y) * [G(x)G(y)]\} = E_x(x, y), \quad (7.46)$$

$$[I(x, y) * G(x)] * G_y(y) = \frac{\partial}{\partial y} \{I(x, y) * [G(x)G(y)]\} = E_y(x, y). \quad (7.47)$$

V uvedených výrazech realizuje první člen v hranatých závorkách projekci. Výstupem dvojrozměrného detektoru hran je vektorové pole $\mathbf{E}(x, y) = (E_x(x, y), E_y(x, y))$. Pro velikost a směr hrany máme

$$|\mathbf{E}(x, y)| = \sqrt{E_x^2(x, y) + E_y^2(x, y)}, \quad \varphi(x, y) = \arctan(E_y(x, y)/E_x(x, y)) + \pi/2. \quad (7.48)$$

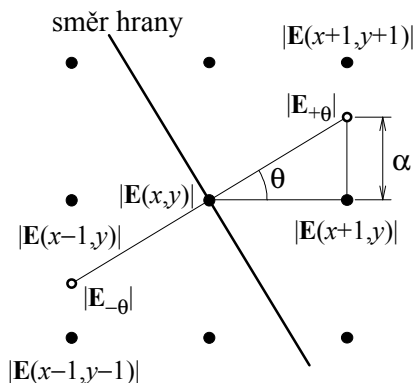
Řekneme, že v nějakém bodě obrazu je hrana, jestliže $|\mathbf{E}(x, y)|$ v tom bodě nabývá lokálního extrému (připomeňme, že jsme se snažili, aby jedné hraně příslušelo pouze jediné maximum) a jestliže je hodnota $|\mathbf{E}(x, y)|$ v tom bodě dostatečně veliká. Předpokládejme nyní, že je obraz popsán diskrétní obrazovou funkcí. Zaměřme se nejdříve na první podmínku. Bod (x, y) může být bodem hrany jen tehdy, jestliže platí (obr. 7.16)

$$|\mathbf{E}_{-\theta}| < |\mathbf{E}(x, y)| > |\mathbf{E}_{+\theta}|. \quad (7.49)$$

Hodnoty $|\mathbf{E}_{+\theta}|$, $|\mathbf{E}_{-\theta}|$ stanovíme interpolací (obr. 7.16). Např. leží-li směr gradientu v prvním oktantu, pak máme

$$|\mathbf{E}_{+\theta}| = \alpha|\mathbf{E}(x+1, y+1)| + (1-\alpha)|\mathbf{E}(x+1, y)|, \quad (7.50)$$

$$|\mathbf{E}_{-\theta}| = \alpha|\mathbf{E}(x-1, y-1)| + (1-\alpha)|\mathbf{E}(x-1, y)|. \quad (7.51)$$



Obr. 7.16: K rozhodnutí maxima

Význam hodnoty α je patrný z obr. 7.16. Podobné vztahy lze snadno odvodit i pro zbývající oktanty. Podmínku, že hodnota $|E(x, y)|$ má být v bodě hrany dostatečně velká, lze realizovat prahováním. Účinné je prahování s hysterezí, při němž se používá dvou hodnot prahu t_{low} , t_{high} . Za body hrany jsou nejprve označeny všechny pixely, kde je velikost hrany větší než t_{high} . Dále jsou pak za body hrany označeny také ty pixely, kde je velikost hrany větší než t_{low} (ale menší než t_{high}) a které sousedí s pixelem, který byl již dříve označen za bod hrany. Tento krok lze opakovat několikrát.

7.2.3 Parametrické modely hrany

Řekněme, že máme rozhodnout, zda je v obraze v místě o souřadnicích (x, y) hrana. V pozitivním případě máme stanovit také její velikost a směr. Můžeme to provést tak, že jistým počtem funkčních hodnot obrazové funkce $f(x, y)$ z okolí bodu (x, y) proložíme vhodnou plochu, která průběh obrazové funkce aproximuje. O existenci hrany pak rozhodneme a její parametry stanovíme na základě vyšetřování průběhu aproximující funkce. Aby se potlačil vliv šumu, prokládá se aproximační funkce zpravidla větším počtem funkčních hodnot obrazové funkce, než je nejnižší teoreticky nutný počet. Proložení lze realizovat minimalizací součtu čtverců diferencí funkčních hodnot aproximační funkce od funkčních hodnot obrazové funkce ve zvolených bodech.

Nejjednodušší plochou, kterou lze jako modelu hrany použít, je rovina $z = ax + by + c$. Nalezneme-li rovinu aproximující průběh obrazové funkce v daném místě, pak snadno získáme velikost hrany $e(x, y) = |\text{grad}(ax + by + c)| = \sqrt{a^2 + b^2}$ a také její směr $\varphi(x, y) = \arctan(b/a) + \pi/2$. Rovinu můžeme nalézt např. tak, aby co nejlépe aproximovala průběh obrazové funkce ve čtyřech bodech $f(x, y)$, $f(x + 1, y)$, $f(x, y + 1)$ a $f(x + 1, y + 1)$. Neznámé hodnoty a , b , c nalezneme minimalizací chyby $\varepsilon = [ax + by + c - f(x, y)]^2 + [a(x + 1) + by + c - f(x + 1, y)]^2 + [ax + b(y + 1) + c - f(x, y + 1)]^2 + [a(x + 1) + b(y + 1) + c - f(x + 1, y + 1)]^2$ uplatněním podmínek

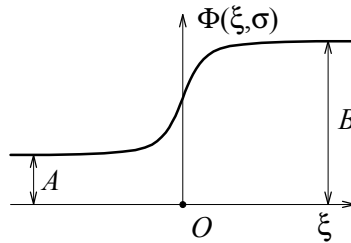
$$\partial\varepsilon/\partial a = 0, \partial\varepsilon/\partial b = 0, \partial\varepsilon/\partial c = 0.$$

Dále ukážeme případ, kdy je jako model hrany použita důmyslnější plocha. Předpokládejme např., že v řezu napříč hranou má jas průběh, který je dobře vystižen funkcí

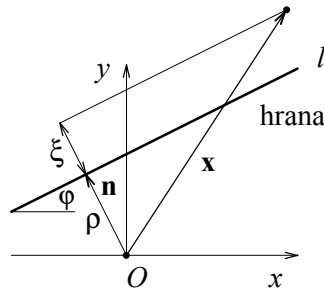
$$B + (A - B)\Phi(\xi, \sigma). \quad (7.52)$$

Význam parametrů A, B objasňuje obr. 7.17. Funkce $\Phi(\xi, \sigma)$ je integrálem gausiánu. Je tedy

$$\Phi(\xi, \sigma) = \int_{t=-\infty}^{\xi} G(t, \sigma) dt, \quad \text{kde} \quad G(t, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{t^2}{2\sigma^2}\right). \quad (7.53)$$



Obr. 7.17: Parametrický model hrany (průběh jasu napříč hranou)



Obr. 7.18: Parametrický model hrany (pohled shora)

Abychom přešli k dvojrozměrnému případu, uvažme, že $\mathbf{n}\mathbf{x} = \rho$, kde $\mathbf{n} = (-\sin\varphi, \cos\varphi)$, $\mathbf{x} = (x, y)$, je rovnice přímky l (obr. 7.18). Jestliže bod reprezentovaný vektorem \mathbf{x} na přímce l neleží, pak vzdálenost bodu od přímky je $\xi = \mathbf{n}\mathbf{x} - \rho = -x\sin\varphi + y\cos\varphi - \rho$. Předpokládáme-li, že hrana leží na přímce l , pak lze průběh jasu v okolí bodu hrany dobře aproximovat funkcí

$$B + (A - B)\Phi(-x\sin\varphi + y\cos\varphi - \rho, \sigma). \quad (7.54)$$

Celkem má tedy popsaný model pět parametrů A , B , σ , φ , ρ . Hodnota $(A - B)$ určuje velikost (výšku) hrany, parametr σ její strmost a parametr φ směr hrany v bodě \mathbf{x} . Parametry lze určit na základě podmínky, aby se v jistém počtu bodů (pixelů) ležících v okolí bodu \mathbf{x} hodnoty funkce (7.54) co nejvíce přimykaly skutečným hodnotám obrazové funkce. Nejmenší počet bodů umožňující úlohu řešit je pět. S ohledem na redukci šumu je však výhodnější volit počet bodů vyšší. Pro hledané parametry lze pak sestavit předeterminovaný systém, který je možné řešit minimalizací chyby. Jistou komplikací ovšem je, že se jedná o problém nelineární. K jeho řešení je proto zapotřebí použít numerických metod.

7.3 Detekce oblastí

V podkapitole 7.1 jsme popsali metody segmentace obrazu, které byly založeny na detekci hranice oblastí. Jiná skupina metod se zaměřuje přímo na vyhledávání celých oblastí v obraze. Oblasti jsou vyhledávány na základě jistým způsobem zvoleného kritéria homogenity oblastí. Takovým kritériem mohou být např. konstantní nebo dosti blízké hodnoty jasu, barvy, pokrytí stejnou texturou atd. Protože ze znalosti hranice dokážeme určit také vnitřek oblastí a naopak, zdá se, že by metody založené na detekci hranic i metody založené na detekci celých oblastí měly být vzájemně rovnocenné. Už na obr. 7.1 jsme však naznačili, že tomu tak v praxi nemusí vždy být. Metody založené na detekci oblastí jsou často upřednostňovány u dosti zašumělých obrazů, kde by detekce hranic byla nespolehlivá. (Šumem zde ale myslíme nejen šum vzniklý např. v kameře; ten zpravidla není rozhodující. Rozhodující bývá nadbytečná a neužitečná, někdy také neočekávaná a obtížně odstranitelná informace v samotné scéně, kterou kamera snímá. Pěkným příkladem může být, když je objekt zájmu částečně překryt např. listím stromu.) Metody detekce oblastí lze rozdělit do tří základních skupin: 1) detekce prahováním, 2) detekce narůstáním oblastí, 3) detekce dělením oblastí. V následujícím textu se zaměříme zejména na prahování, a to proto, že ostatní uvedené techniky jsou často konstruovány heuristicky, a přesahují tedy rámec tohoto textu.

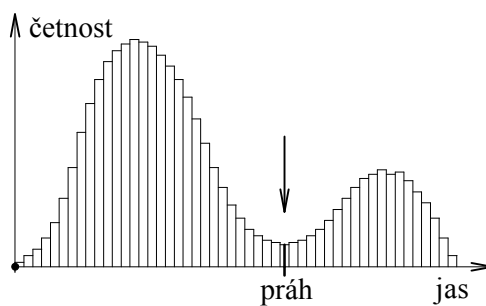
7.3.1 Prahování

Prahování je nejjednodušší metodou detekce celých oblastí v obraze. V nekomplikovaných případech se při tom ale současně jedná o metodu rychlou a spolehlivou. Výsledkem prahování je binární obraz, v němž je bodům (pixelům) nalezených oblastí obvykle přiřazena hodnota 1, zatímco zbývajícím bodům (bodům pozadí) je přiřazena hodnota 0. Při prahování se vychází z předpokladu, že body hledaných oblastí mají stejný nebo dosti podobný jas. Jedna z možných realizací prahování spočívá v tom, že pixel je označen jako bod oblasti, jestliže jeho jas padne do intervalu $\langle a, b \rangle$. Velmi často je v rozhodovacím kritériu použita pouze jediná hodnota t - tzv. práh. Body, jejichž jas je vyšší než hodnota práhu, jsou detekovány jako body hleda-

ných oblastí a body, jejichž jas je nižší, jsou detekovány jako body pozadí (případně naopak). *Prahování* se tedy děje na základě předpisu

$$g(x, y) = \begin{cases} 1, & f(x, y) \in \langle a, b \rangle \\ 0, & \text{jinak} \end{cases} \quad \text{nebo} \quad g(x, y) = \begin{cases} 1, & f(x, y) \geq t \\ 0, & \text{jinak} \end{cases} . \quad (7.55)$$

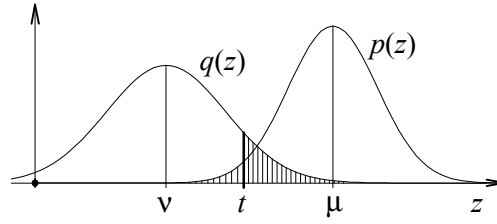
Úspěšnost prahování závisí na znalosti správné hodnoty prahu. Jestliže tuto hodnotu neznáme, je možné pokusit se ji stanovit na základě informací získaných z obrazu, který má být segmentován. Pro obrazy s bimodálním histogramem (histogramem se dvěma vrcholy) jasu se např. často doporučuje volit jako hodnotu t prahu hodnotu, v níž histogram dosahuje mezi oběma vrcholy minima (obr. 7.19). Lze snadno domyslet, že tato heuristická metoda má racionální základ. Předpokládá totiž, že v obraze existují dva druhy pixelů: pixely náležící hledaným oblastem a pixely náležící pozadí. Oba druhy pixelů jsou při tom relativně četné a mají dosti odlišný jas, což v histogramu dává vzniknout zmíněným dvěma vrcholům. Nejsou-li na druhé straně uvedené předpoklady splněny, nemusí být výsledky metody uspokojivé. Ne vždy je také *histogram bimodální*.



Obr. 7.19: Bimodální histogram jasu

V případech, kdy je obraz sice kontrastní, avšak v různých svých částech má nerovnoměrnou úroveň jasu, nelze často najít jedinou hodnotu prahu tak, aby vyhovovala pro všechny části obrazu. V takovém případě lze použít prahování s proměnnou hodnotou prahu, kdy se pro různá místa obrazu používá různých hodnot prahu. Metodu lze realizovat tak, že se obraz rozdělí na několik částí a pro každou se najde individuální hodnotu prahu. Je-li histogram úrovní jasů vyšetřované části obrazu bimodální, lze hodnotu prahu stanovit např. postupem popsáním v předchozím odstavci. Jestliže ve vyšetřované části obrazu velmi převažují buď body oblasti nebo body pozadí, pak hodnotu prahu z histogramu uvažované části obrazu spolehlivě stanovit nelze. V takovém případě lze hodnotu prahu stanovit např. jako průměr hodnot prahů z přílehlých oblastí.

Ukážeme dále důmyslnější metodu stanovení prahu, která je založena na minimalizaci chyby a na využití počtu pravděpodobnosti. Metoda je pěkná. Poukazuje na skutečnost, že, přesně vzato, někdy ani nemusí být možné pro každý pixel správně



Obr. 7.20: Stanovení prahu minimalizací chyby

rozlišit, zda se jedná o pixel objektu nebo o pixel pozadí. Metoda minimalizace chyby je typickým postupem, který lze v dané situaci uplatnit. Předpokládejme, že jasy pixelů oblastí, které mají být detekovány, mají normální rozložení hustoty pravděpodobnosti $p(z)$ se střední hodnotou μ a směrodatnou odchylkou σ (obr. 7.20). Podobně předpokládáme, že také jasy pixelů pozadí mají normální rozložení hustoty pravděpodobnosti $q(z)$ se střední hodnotou ν a směrodatnou odchylkou τ (přísně teoreticky vzato nemohou být ovšem rozložení normální, protože prakticky jsou jasové úrovně nenulové pouze v jistém intervalu; může se však jednat o dobré přiblížení). Dále předpokládejme, že podíl pixelů hledaných oblastí na celkovém počtu pixelů obrazu je θ a že platí $\mu > \nu$. Předpokládejme nyní, že zvolíme práh t . Označme $P(t)$ pravděpodobnost jevu, že bod hledané oblasti bude nesprávně vyhodnocen jako bod pozadí. Dále označme $Q(t)$ pravděpodobnost jevu, že bod pozadí bude správně vyhodnocen jako bod pozadí. Pak $1 - Q(t)$ je pravděpodobnost, že bod pozadí bude nesprávně vyhodnocen jako bod oblasti. Pro pravděpodobnosti $P(t)$ a $Q(t)$ máme (obr. 7.20)

$$P(t) = \int_{-\infty}^t p(z) dz, \quad Q(t) = \int_{-\infty}^t q(z) dz. \quad (7.56)$$

Dále musíme uvážit, že poměrné zastoupení bodů objektu v obraze je θ a poměrné zastoupení bodů pozadí v obraze je $(1 - \theta)$. Výsledná pravděpodobnost chybné klasifikace je pak

$$\varepsilon = \theta P(t) + (1 - \theta)[1 - Q(t)]. \quad (7.57)$$

K nalezení minima výraz (7.57) derivujeme a derivaci položíme rovnu nule. Dostaneme

$$\frac{\partial \varepsilon}{\partial t} = \theta \frac{\partial P(t)}{\partial t} - (1 - \theta) \frac{\partial Q(t)}{\partial t} = \theta p(t) - (1 - \theta)q(t) = 0. \quad (7.58)$$

Odtud máme

$$(1 - \theta)q(t) = \theta p(t). \quad (7.59)$$

Podle předpokladu mají hustoty pravděpodobnosti $p(z)$ i $q(z)$ normální rozložení. Je tedy

$$p(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(t - \mu)^2}{2\sigma^2}\right], \quad q(t) = \frac{1}{\tau\sqrt{2\pi}} \exp\left[-\frac{(t - \nu)^2}{2\tau^2}\right]. \quad (7.60)$$

Dosazením vztahů (7.60) do (7.59) a logaritmováním máme

$$\ln(1 - \theta) - \ln \tau - \frac{(t - \nu)^2}{2\tau^2} = \ln \theta - \ln \sigma - \frac{(t - \mu)^2}{2\sigma^2}. \quad (7.61)$$

Po úpravě dostaneme

$$\tau^2(t - \mu)^2 - \sigma^2(t - \nu)^2 = 2\sigma^2\tau^2 \ln \frac{\tau\theta}{\sigma(1 - \theta)}. \quad (7.62)$$

Výraz (7.62) je kvadratickou rovnicí pro t . K řešení je zapotřebí znát hodnoty μ , σ , ν , τ , θ . Povšimněme si blíže ještě případu, kdy $\theta = 1/2$ a $\sigma = \tau$. V tomto případě se rovnice (7.62) zjednoduší na tvar

$$(t - \mu)^2 = (t - \nu)^2. \quad (7.63)$$

Řešením rovnice (7.63) je intuitivně očekávaná hodnota $t = (\mu + \nu)/2$.

Třebaže jsme metodu v úvodu označili za pěknou, pozorný čtenář nyní již pravděpodobně vidí určité nedostatky. Metoda předpokládá, že hustoty $p(z)$ a $q(z)$ zůstávají nezměněné pro všechny zpracovávané obrazy, na než je pak aplikována stanovená hodnota t prahu. To je v praxi ne vždy splnitelné, protože světelné podmínky při snímání jednotlivých obrazů se mohou měnit. Po pravdě je zapotřebí připustit, že v praxi bychom hodnotu prahu velmi často určili zkusmo.

7.4 Segmentace jako úloha o minimalizaci funkcionálu

V této podkapitole se budeme zabývat důmyslnými a náročnějšími matematickými metodami segmentace. Necht Ω je oblast, nad níž je definován obraz. Dále necht $u_0(x)$ je obraz, který má být podroben segmentaci. Bez újmy na obecnosti můžeme předpokládat $0 \leq u_0(x) \leq 1$, $x \in \Omega$ (předpokládáme obrazy ve stupních šedi). Hledáme obraz $u(x)$, který je opět definován nad Ω . V hledaném obraze je $K \subset \Omega$ uzavřená množina bodů, v níž může být hledaný obraz nespojitý. Je-li N topologická dimenze množiny Ω , pak topologická dimenze množiny K je $N - 1$. Přesně vzato, spočívá řešení úlohy o segmentaci obrazu v nalezení dvojice (u, K) tak, aby byly splněny jisté vlastnosti, které uvedeme v dalším textu.

Mumford a Shah navrhli řešit úlohu o segmentaci obrazu minimalizací funkcionálu

$$F(u, K) = \int_{\Omega \setminus K} (u - u_0)^2 dx + \alpha \int_{\Omega \setminus K} |\nabla u|^2 dx + \beta \int_K d\sigma. \quad (7.64)$$

Ve funkcionálu (7.64) mají jednotlivé členy následující význam: α a β jsou předem zvolené konstanty - váhy, které bývají stanoveny experimentálně. První člen funkcionálu zajišťuje, aby se nalezený obraz u co nejvíce přimyká obrazu zadanému u_0 .

Druhý člen vyjadřuje očekávání, že mimo body patřící do K (v nichž se připouští nespojitost) by měl být nalezený obraz u co nejhladší. Třetí člen vyjadřuje požadavek, aby délka (míra) oblasti, v níž nalezený obraz nespojitý, byla co nejmenší.

Někteří autoři později kritizovali, že třetí člen ve funkcionálu (7.64) není definován dostatečně přesně a navrhli funkcionál ve tvaru

$$F(u, K) = \int_{\Omega \setminus K} (u - u_0)^2 dx + \alpha \int_{\Omega \setminus K} |\nabla u|^2 dx + \beta \mathcal{H}^{N-1}(K), \quad (7.65)$$

kde $\mathcal{H}^{N-1}(K)$ je Hausdorffova míra množiny K . Zjednodušením funkcionálu (7.65) pak dále můžeme získat funkcionál

$$E(u, K) = \int_{\Omega \setminus K} (u - u_0)^2 dx + \beta \mathcal{H}^{N-1}(K). \quad (7.66)$$

Ve srovnání s funkcionálem (7.65) chybí ve funkcionálu (7.66) člen požadující hladkost hledaného obrazu. To má své opodstatnění. Funkcionálu lze dobře využít tehdy, jestliže předpokládáme, že je jas hledaného obrazu nad jednotlivými oblastmi vymezenými množinou K konstantní. Druhý člen ve funkcionálu (7.65) pak vymizí.

Minimalizace funkcionálů (7.64) až (7.66) je obtížná. Obtížnost spočívá v tom, že se minimalizuje vzhledem ke dvěma neznámým, které mají rozdílnou povahu: u je funkce nad N -dimenzionálním prostorem; K je $(N - 1)$ dimenzionální množina bodů. Určitého zjednodušení lze dosáhnout, jestliže se předpokládá, že neznámou je jen obraz u a že K lze odvodit z u jako množinu bodů, kde u vykazuje skok. Označme tuto množinu S_u . Funkcionál (7.64) lze pak zapsat ve tvaru

$$G(u) = \int_{\Omega} (u - u_0)^2 dx + \alpha \int_{\Omega} |\nabla u|^2 dx + \beta \mathcal{H}^{N-1}(S_u). \quad (7.67)$$

7.4.1 Aktivní kontury „bez hran“

Tato metoda spadá do segmentace pomocí minimalizace funkcionálu. Úloha je řešena pomocí tzv. level setů. Zajímavá je myšlenka nevyužít přímo hranových bodů v obraze (proto se v názvu metody objevuje termín *bez hran*). V původní variantě metody se uvažují objekty, o nichž se předpokládá, že mají všechny stejný a konstantní jas, a dále pozadí, u něhož se také předpokládá konstantní jas. Funkcionál, který se má minimalizovat (Chan-Vese funkcionál, CV funkcionál), je tvaru

$$\begin{aligned} F(C, c_1, c_2) = & \mu \cdot \text{length}(C) + \nu \cdot \text{area}(\text{inside}(C)) + \\ & + \lambda_1 \int_{\text{inside}(C)} |u_0 - c_1|^2 dx dy + \\ & + \lambda_2 \int_{\text{outside}(C)} |u_0 - c_2|^2 dx dy, \end{aligned} \quad (7.68)$$

kde C , c_1 , c_2 jsou neznámé (hranice objektů, jas objektů, jas pozadí); $\mu \geq 0$, $\nu \geq 0$, $\lambda_1, \lambda_2 \geq 0$ jsou předem zvolené parametry (obvykle se ale volí $\gamma = 0$, $\lambda_1 = \lambda_2$). Lze si povšimnout podobnosti CV funkcionalu s MS funkcionalem, který je tvaru

$$F^{MS}(u, C) = \int_{\Omega \setminus C} (\alpha |\nabla u|^2 + \beta |u - u_0|^2) dx dy + \text{length}(C). \quad (7.69)$$

Funkcional CV je skutečně speciálním případem funkcionalu MS, když funkci jasu nad objekty (pozadím) předpokládáme konstantní (hodnoty c_1, c_2). Zavedeme takovou funkci ϕ , že na hranici objektů je $\phi = 0$, uvnitř objektů je $\phi > 0$ a vně objektů (v místech pozadí) je $\phi < 0$ (level set funkce). K dalšímu odvození využijeme Heavisidovu funkci

$$H(x) = \begin{cases} 1, & \text{když } x \geq 0, \\ 0, & \text{když } x < 0. \end{cases} \quad (7.70)$$

Poznamenejme dále, že platí $\frac{d}{dx} H(x) = \delta(x)$. S využitím doposud uvedeného lze funkcional (7.68) přepsat ve tvaru

$$\begin{aligned} F(\phi, c_1, c_2) = & \mu \int_{\Omega} \delta(\phi) |\nabla \phi| dx dy + \nu \int_{\Omega} H(\phi) dx dy + \\ & + \lambda_1 \int_{\Omega} |u_0 - c_1|^2 H(\phi) dx dy + \\ & + \lambda_2 \int_{\Omega} (u_0 - c_2) (1 - H(\phi)) dx dy. \end{aligned} \quad (7.71)$$

Jestliže ϕ ponecháme konstantní a minimalizujeme $F(\phi, c_1, c_2)$ s ohledem na c_1, c_2 , dostaneme

$$c_1(\phi) = \frac{\int_{\Omega} u_0 H(\phi) dx dy}{\int_{\Omega} H(\phi) dx dy}, \quad (7.72)$$

$$c_2(\phi) = \frac{\int_{\Omega} u_0 (1 - H(\phi)) dx dy}{\int_{\Omega} (1 - H(\phi)) dx dy}. \quad (7.73)$$

Je zřejmé, že hodnoty $c_1(\phi)$, $c_2(\phi)$ mají význam střední hodnoty jasu nad objekty, případně nad pozadím. Nyní naopak minimalizujme $F(\phi, c_1, c_2)$ podle ϕ . Obdržíme následující Euler-Lagrangeovu rovnici

$$\begin{cases} \frac{\partial \phi}{\partial t} = \delta(\phi) \left[\mu \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) - \nu - \lambda_1 (u_0 - c_1)^2 + \lambda_2 (u_0 - c_2)^2 \right] \text{ in } \Omega, \\ \frac{\delta(\phi)}{\nabla \phi} \frac{\partial \phi}{\partial \omega} = 0 \text{ na } \partial \Omega. \end{cases} \quad (7.74)$$

Aby bylo možné rovnici numericky řešit, zavádí se regularizace odstraňující nespojitost funkce $H(x)$. Funkci $H(x)$ lze nahradit přibližnou hodnotou $H_{\varepsilon}(x)$ dle vztahu

$$H_{\varepsilon}(x) = \frac{1}{2} \left(1 + \frac{2}{\pi} \arctan \left(\frac{x}{\varepsilon} \right) \right), \quad \delta_{\rho}(x) = H'_{\varepsilon}(x) = \frac{1}{\pi} \cdot \frac{\varepsilon}{\varepsilon^2 + x^2}. \quad (7.75)$$

Nyní již lze rovnici (7.74) řešit postupy, které jsou v daném kontextu obvyklé. Jedná se vlastně o obyčejnou integraci. Jako počáteční hodnoty funkce $\phi(x)$ se berou znaménkem opatřené vzdálenosti jednotlivých bodů obrazu od počáteční předem zvolené (odhadnuté) křivky. Funkce ϕ , která je řešením nalezena, ukazuje svými kladnými, zápornými a nulovými hodnotami, kde se nachází objekty, pozadí a hranice mezi nimi. K dosažení dobrých výsledků (numerické stability) je zapotřebí dbát určitých pravidel (vhodná integrační schémata). Čas od času může být také zapotřebí hodnoty funkce $\phi(x)$ přepočítat pomocí tzv. reinicializace. Rozbor podrobností ovšem přesahuje možnosti tohoto stručného úvodního textu.

7.5 Zpracování binárních obrazů

Binárním obrazem nazýváme takový obraz, v němž obrazová funkce v každém bodě (v každém pixelu) nabývá jedné ze dvou možných hodnot. Binární obrazy jsou zpravidla výsledkem metod provádějících segmentaci obrazu (v pixelech náležících objektům např. obrazová funkce nabývá hodnoty 1, v pixelech pozadí nabývá hodnoty 0). Před tím, než jsou binární obrazy analyzovány, lze je zpracovat některým ze speciálních postupů, které dále ve stručném přehledu popíšeme.

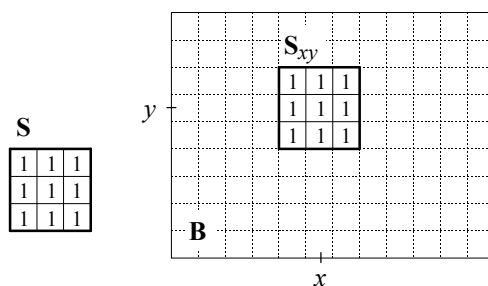
7.5.1 Matematická morfologie

Tvůrcem matematické morfologie je Serra [15]. Jako příklad o něco mladší publikace zabývající se touto problematikou uveďme alespoň práci [7]. Teorie matematické morfologie je dosti obsáhlá, a proto můžeme v tomto textu uvést pouze základní informace. Omezíme se pouze na binární obrazy. Označme \mathbf{B} vstupní binární obraz. Dále budeme pracovat s pomocným binárním obrazem \mathbf{S} . Význam tohoto pomocného obrazu bude obdobný jako význam masky u konvoluce - budeme jej postupně přikládat na různá místa obrazu \mathbf{B} . Označení $\mathbf{S}_{x,y}$ budeme používat pro obraz, který vznikne translací pomocného obrazu \mathbf{S} tak, aby počátek obrazu \mathbf{S} padl do bodu o souřadnicích (x, y) (obr. 7.21). Základními operacemi matematické morfologie jsou *eroze* a *dilatace*. Tyto operace jsou definovány následujícími vztahy (obr. 7.22)

$$\mathbf{E} = \mathbf{B} \otimes \mathbf{S} = \{(x, y) | \mathbf{S}_{x,y} \subseteq \mathbf{B}\}, \quad (7.76)$$

$$\mathbf{D} = \mathbf{B} \oplus \mathbf{S} = \{(x, y) | \mathbf{S}_{x,y} \cap \mathbf{B} \neq \emptyset\}. \quad (7.77)$$

Vysvětleme podrobněji význam vztahu (7.76). Erozí binárního obrazu \mathbf{B} za použití masky \mathbf{S} vznikne obraz \mathbf{E} , který je opět binární. Předpokládejme, že jednotlivé body binárního obrazu nesou hodnotu 0 nebo 1. V bodě o souřadnicích (x, y) je v obraze \mathbf{E} hodnota 1, jestliže je v obraze \mathbf{B} hodnota 1 alespoň na těch místech, kde je hodnota 1 v masce $\mathbf{S}_{x,y}$. Jinak je v obraze \mathbf{E} v bodě o souřadnicích (x, y) hodnota 0. Ve vztahu (7.76) jsou obrazy formálně reprezentovány jako množiny - jedná se o množiny pixelů nesoucích hodnotu 1. Analogicky by bylo možné interpretovat také vztah (7.77).



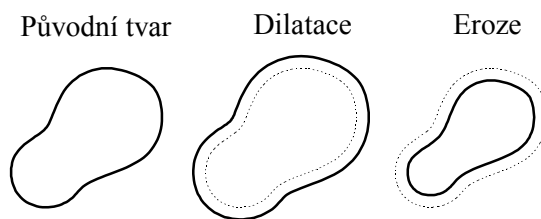
Obr. 7.21: K významu masky u morfologických operací

Dalšími často používanými operacemi matematické morfologie jsou *otevření* a *uzávěr*. Otevření je definováno jako eroze následovaná dilatací (vztah (7.78)). Uzavření je naopak dilatace následovaná erozí (vztah (7.79)). Máme tedy

$$\mathbf{B} \circ \mathbf{S} = (\mathbf{B} \otimes \mathbf{S}) \oplus \mathbf{S}, \quad (7.78)$$

$$\mathbf{B} \bullet \mathbf{S} = (\mathbf{B} \oplus \mathbf{S}) \times \mathbf{S}. \quad (7.79)$$

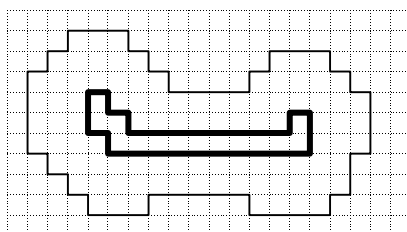
Z definice je zřejmé, že otevření eliminuje malé a tenké objekty a rozděljuje objekty v místech, kde jsou tenké. Uzavření naopak vyplňuje malé a tenké díry v objektech a spojuje objekty, které leží blízko sebe. Poznamenejme, že někdy může být výhodné použít většího počtu otevření nebo uzavření po sobě. Alternativně může také být otevření a uzavření realizováno také tak, že je použito většího počtu operací eroze a stejného počtu operací dilatace.



Obr. 7.22: Dilatace a eroze

Užitečnou úpravou binárního obrazu, při níž může být použito morfologických operací, je ztenčování. Cílem ztenčování je reprezentovat objekty v obraze jako lineární útvary (obr. 7.23). Ztenčování může být realizováno pomocí opakovaně prováděné eroze (tj. opakovaným odstraňováním krajních pixelů z objektu). Postup provádění každého erozního kroku je při tom modifikován tak, aby nedošlo k porušení souvislosti objektu. Při praktickém výpočtu jsou dle vztahu (7.76) nejprve detekovány pixely objektu kandidující na to, že do nich bude zapsána hodnota 0 (tj. budou z objektu odstraněny). Tato hodnota je pak ale skutečně zapsána jen do

některých z kandidujících pixelů. Pixely jsou vybrány tak, aby zápisem hodnoty 0 nedošlo k rozdělení zbývající části objektu na více částí. Proces opakovaného provádění eroze pokračuje tak dlouho, dokud zbylá část objektu nemá všude tloušťku jeden pixel.



Obr. 7.23: Objekt před a po ztenčení

Kapitola 8

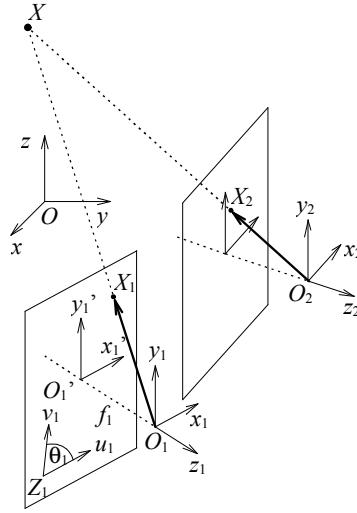
Zpětná stereoprojekce

Máme-li k dispozici několik obrazů téže scény, které byly získány několika kamerami umístěnými ve scéně nebo jedinou kamerou umístěnou postupně do různých míst, pak lze za jistých okolností provést rekonstrukci scény. Nejmenší počet kamer (případně míst), které jsou k provedení rekonstrukce zapotřebí, jsou dvě (obr. 8.1). Rekonstrukcí při tom rozumíme, že pro zvolené body dokážeme na základě souřadnic jejich obrazů zachycených kamerami vypočítat trojrozměrné souřadnice ve scéně. Řešení naznačeného problému podrobněji popíšeme v následujících podkapitolách.

8.1 Model kamery

Předpokládáme, že vztah mezi objekty pozorovanými kamerou a jejich obrazy je lineární projektivní transformací. Modelem kamery je tedy dírková komora. Ve scéně zavedeme globální souřadný systém (O, x, y, z) (obr. 8.1). Uvažujme i -tou kameru (případně i -tou polohu kamery). Pro označení veličin vztahujících se k i -té kameře použijeme symbolů s indexem i . Střed projekce i -té kamery je O_i ; π_i je zobrazovací rovina. Kolmice vedená z O_i na rovinu π_i je optická osa kamery. V zobrazovací rovině kamery leží zařízení zachycující obraz (jedná se například o pole snímacích elementů CCD čipu, případně o světlocitlivou vrstvu). Souřadnice bodů v obrazech měříme v souřadném systému (Z_i, u_i, v_i) (obr. 8.1). V případě CCD kamery jsou často souřadné osy tohoto systému hranami pole snímacích elementů. Ortogonální souřadný systém (O_i, x_i, y_i, z_i) je souřadným systémem kamery. Jeho osa z_i je identická s optickou osou kamery, osa x_i je rovnoběžná s osou u_i (v případě CCD kamery tedy s hranou pole snímacích elementů). Dále v obraze zavedeme souřadný systém (O'_i, x'_i, y'_i) . Bod O'_i leží v místě, kde optická osa kamery protíná zobrazovací rovinu. Osa x'_i je rovnoběžná s osou x_i , osa y'_i je rovnoběžná s osou y_i . Vzdálenost $f_i = \text{dist}(O_i, O'_i)$ je ohniskovou vzdáleností kamery.

Uvažujme nějaký bod X scény. V souřadném systému (O, x, y, z) je bod X reprezentován vektorem $\mathbf{x} = (x, y, z)^\top$. V souřadném systému (O_i, x_i, y_i, z_i) je tentýž bod reprezentován vektorem $\mathbf{x}_i = (x_i, y_i, z_i)^\top$. Pro transformaci ze souřadného systému (O_i, x_i, y_i, z_i) kamery do globálního souřadného systému (O, x, y, z) scény můžeme



Obr. 8.1: Zpětná stereoprojekce

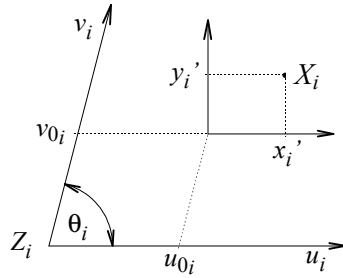
psát

$$\mathbf{x} = \mathbf{R}_i \mathbf{x}_i + \mathbf{o}_i, \quad (8.1)$$

kde matice \mathbf{R}_i popisuje rotaci kamery a vektor \mathbf{o}_i (zadaný v souřadném systému (O, x, y, z)) reprezentuje střed projekce O_i . Nechť X_i je bod v zobrazovací rovině i -té kamery a tento bod nechť je obrazem bodu X scény. Souřadnice bodu X_i měřené v souřadném systému (Z_i, u_i, v_i) jsou (u_i, v_i) . Tyto souřadnice uspořádáme do trojrozměrného vektoru $\mathbf{u}_i = (u_i, v_i, 1)^\top$. V souřadném systému (O'_i, x'_i, y'_i) je poloha bodu X_i popsána vektorem $\mathbf{x}'_i = (x'_i, y'_i, 1)^\top = \mathbf{Q}_i \mathbf{u}_i$, kde \mathbf{Q}_i je matice popisující transformaci ze souřadného systému (Z_i, u_i, v_i) do souřadného systému (O'_i, x'_i, y'_i) . Snadno ověříme, že platí

$$\begin{aligned} \mathbf{Q}_i &= \left(\begin{bmatrix} 1 & 0 & u_{0i} \\ 0 & 1 & v_{0i} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & s_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\cotg \theta_i & 0 \\ 0 & 1/\sin \theta_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)^{-1} \\ &= \begin{bmatrix} 1 & -\cotg \theta_i & u_{0i} \\ 0 & s_i / \sin \theta_i & v_{0i} \\ 0 & 0 & 1 \end{bmatrix}^{-1}. \end{aligned} \quad (8.2)$$

Jednotlivé symboly v matici \mathbf{Q}_i mají následující význam (obr. 8.2). Hodnoty u_{0i} , v_{0i} jsou souřadnice (měřené v souřadném systému (Z_i, u_i, v_i)) bodu, v němž optická osa protíná zobrazovací rovinu kamery. Úhel θ_i modeluje možnou odchylku pole snímacích elementů CCD senzoru od ortogonality, případně možnou úchylku kolmosti zobrazovací roviny a optické osy. Souřadný systém (Z_i, u_i, v_i) tedy nemusí být ortogonální, ale v praxi je hodnota θ_i obvykle velmi blízká $\pi/2$. Význam měřítkového faktoru s_i je následující: U kamer s CCD čipem zohledňuje možné difference rozměrů

Obr. 8.2: K významu prvků matice \mathbf{Q}_i .

snímacích elementů ve směru osy u a v . Jestliže je výstup kamery analogový, pak parametr s_i zohledňuje také možné zkreslení rozměrů snímacích elementů způsobené následnou zpětnou digitalizací analogového signálu. Ze vztahu (8.2) je zřejmé, že parametrem s_i korigujeme pouze poměrné zkreslení délek na osách - skutečná celková velikost obrazu bude vzata v úvahu hodnotou ohniskové vzdálenosti. Poznamenejme, že matice \mathbf{Q}_i^{-1} popisuje transformaci v opačném směru než matice \mathbf{Q}_i , tedy ze souřadného systému (O'_i, x'_i, y'_i) do souřadného systému (Z_i, u_i, v_i) . Zdá se názornější sestavit právě matici \mathbf{Q}_i^{-1} . Matici \mathbf{Q}_i získáme inverzí. Tento postup jsme uplatnili ve vztahu (8.2). Zavedeme dále matici

$$\mathbf{F}_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -f_i \end{bmatrix}. \quad (8.3)$$

Snadno ověříme, že platí

$$\mathbf{x}_i = \lambda_i \mathbf{F}_i \mathbf{Q}_i \mathbf{u}_i, \quad (8.4)$$

kde λ_i je reálný parametr. S využitím vztahu (8.1) obdržíme

$$\mathbf{x} = \lambda_i \mathbf{R}_i \mathbf{F}_i \mathbf{Q}_i \mathbf{u}_i + \mathbf{o}_i, \quad (8.5)$$

případně

$$\lambda_i \mathbf{u}_i = (\mathbf{R}_i \mathbf{F}_i \mathbf{Q}_i)^{-1} (\mathbf{x} - \mathbf{o}_i). \quad (8.6)$$

Parametry obsažené v maticích \mathbf{F}_i , \mathbf{Q}_i ($f_i, u_{0i}, v_{0i}, \theta_i, s_i$) jsou vnitřní (tzv. intrinsické) parametry kamery. Matice \mathbf{Q}_i zohledňuje některé rozdíly mezi skutečnou a ideální kamerou. V jednoduchých teoretických úvahách se někdy předpokládá $\mathbf{Q}_i = \mathbf{I}$ (jednotková matice). Hodnoty \mathbf{R}_i , \mathbf{o}_i popisující polohu kamer v prostoru se nazývají vnějšími (tzv. extrinsickými) parametry kamery. Základní úlohou počítačového stereovidění je na základě souřadnic (u_i, v_i) obrazů nějakého bodu scény zjistit jeho prostorové souřadnice (x, y, z) . K řešení této úlohy je zpravidla nejprve nutné provést tzv. kalibraci kamery. Úkolem kalibrace je stanovit vnitřní a vnější parametry kamery. Jen ve vyjíměčných případech, kdy jsou parametry kamery předem známy,

může kalibrace chybět (od výrobce např. můžeme znát hodnoty vnitřních parametrů, umístění kamer v prostoru můžeme někdy dostatečně přesně změřit).

Vztahy popisující promítání realizované kamerou se zjednoduší, jestliže použijeme homogenních souřadnic. (Předpokládáme zde, že je čtenář obeznámen alespoň s elementárními pojmy z projektivní geometrie. V případě potřeby lze potřebné znalosti získat např. z prací [17], [3], [13]). Ostatně již dříve jsme vektor \mathbf{u}_i zavedli jako $\mathbf{u}_i = (u_i, v_i, 1)^\top$, což lze interpretovat jako homogenní souřadnice bodu X_i v dvojrozměrném projektivním prostoru. Dále navíc připustíme, aby homogenní souřadnice nabývala obecné hodnoty w_i , tedy $\mathbf{u}_i = (w_i u_i, w_i v_i, w_i)^\top$. Podobně popíšeme v homogenních souřadnicích i polohu bodu X vzhledem ke globálnímu souřadnému systému scény. Položíme $\mathbf{x} = (x, y, z, 1)^\top$. Projektivní transformaci realizovanou kamerou lze nyní zapsat ve tvaru

$$\mathbf{u}_i = \mathbf{P}_i \mathbf{x}. \quad (8.7)$$

Matice \mathbf{P}_i rozměru 3×4 popisuje transformaci. Vztahy (8.5), (8.6), (8.7) popisují různými prostředky tutéž transformaci a jsou vzájemně ekvivalentní. V dalším textu budeme podle potřeby používat vždy ten tvar, který je pro danou úlohu nejvýhodnější. Porovnáním vztahu (8.7) se vztahem (8.6) zjišťujeme, že matice \mathbf{P}_i zahrnuje vliv vnějších i vnitřních parametrů kamery.

8.2 Dvojice kamer s rovnoběžnými optickými osami

Nejprve ukážeme jednoduchý případ ilustrující, že rekonstrukce je možná. Předpokládejme, že máme dvě kamery, které jsou uspořádány podle následujících podmínek (obr. 8.3): 1) Obě kamery mají rovnoběžné optické osy. 2) Předpokládáme, že platí $\mathbf{Q}_1 = \mathbf{Q}_2 = \mathbf{I}$, tedy $x'_1 = u_1$, $y'_1 = v_1$, $x'_2 = u_2$, $y'_2 = v_2$. 3) Souřadné osy u_1 , u_2 leží v jedné přímce. 4) Souřadné osy v_1 , v_2 jsou rovnoběžné. 5) Vzdálenost optických středů kamer je b . 6) Počátek globálního souřadného systému scény pólí spojnicí O_1 , O_2 a osy globálního systému jsou rovnoběžné s osami $u_1(u_2)$, $v_1(v_2)$. 7) $f_1 = f_2 = f$. Uvažujme ve scéně bod X o souřadnicích (x, y, z) . Bod X se promítá do bodů X_1 , X_2 ležících v zobrazovacích rovinách první a druhé kamery. Souřadnice bodů X_1 , X_2 měřené v obrazech získaných kamerami jsou (u_1, v_1) , (u_2, v_2) . Z podobnosti trojúhelníků snadno obdržíme následující vztahy (předpokládáme, že ohnisková vzdálenost f se zadává jako kladná hodnota)

$$\frac{x + b/2}{-z} = \frac{u_1}{f}, \quad \frac{x - b/2}{-z} = \frac{u_2}{f}. \quad (8.8)$$

Analogické vztahy snadno zapíšeme také pro souřadnici y . Máme

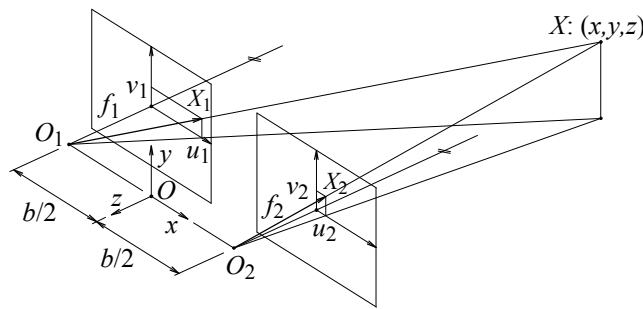
$$\frac{y}{-z} = \frac{v_1}{f}, \quad \frac{y}{-z} = \frac{v_2}{f}. \quad (8.9)$$

V rovnicích (8.8) jsou x , z neznámými. Ostatní hodnoty jsou známé. Řešením soustavy pro z získáme

$$z = -\frac{bf}{u_1 - u_2}. \quad (8.10)$$

S využitím rovnic (8.8), (8.9) pak snadno také nalezneme hodnoty x , y . Vychází

$$x = \frac{b(u_1 + u_2)}{2(u_1 - u_2)}, \quad y = \frac{b(v_1 + v_2)}{2(u_1 - u_2)}. \quad (8.11)$$



Obr. 8.3: Kamery s rovnoběžnými optickými osami

8.3 Absolutní kalibrace a rekonstrukce

V tomto odstavci ukážeme postup, v němž se kalibrace provádí na základě toho, že pro jistý počet tzv. kalibračních bodů známe jak souřadnice u_i , v_i jejich obrazů tak i jejich souřadnice x , y , z v souřadném systému scény (název absolutní kalibrace je odvozen od požadavku znát souřadnice x , y , z kalibračních bodů). Zaměříme se na případ, kdy máme pouze dvě kamery ($i = 1, 2$). K odvození výsledných vztahů použijeme homogenních souřadnic a vyjdeme ze vztahu (8.7). Matici \mathbf{P}_i rozepíšeme do tří řádkových vektorů rozměru 1×4 . Máme pak $\mathbf{P}_i = (\mathbf{p}_{i,1}, \mathbf{p}_{i,2}, \mathbf{p}_{i,3})^\top$. Rovnici (8.7) lze nyní přepsat takto

$$\begin{bmatrix} w_i u_i \\ w_i v_i \\ w_i \end{bmatrix} = \begin{bmatrix} \mathbf{p}_{i,1} \\ \mathbf{p}_{i,2} \\ \mathbf{p}_{i,3} \end{bmatrix} \mathbf{x}. \quad (8.12)$$

Předpokládejme nyní, že známe projekční matice \mathbf{P}_1 , \mathbf{P}_2 obou kamer. Dále předpokládejme, že je ve scéně bod, jehož obraz získaný první kamerou má souřadnice u_1 , v_1 a obraz získaný druhou kamerou má souřadnice u_2 , v_2 . Ukážeme, jak lze rekonstruovat trojrozměrné souřadnice x , y , z tohoto bodu. Položme $\mathbf{x} = (x, y, z, 1)$. Z posledního řádku rovnice (8.12) máme

$$w_1 = \mathbf{p}_{1,3}\mathbf{x}, \quad w_2 = \mathbf{p}_{2,3}\mathbf{x}. \quad (8.13)$$

Dosažením rovnic (8.13) do prvních dvou řádků vztahu (8.12) rozepsaného pro první i druhou kameru dostaneme

$$\begin{bmatrix} \mathbf{p}_{1,1}\mathbf{x} - u_1\mathbf{p}_{1,3}\mathbf{x} \\ \mathbf{p}_{1,2}\mathbf{x} - v_1\mathbf{p}_{1,3}\mathbf{x} \\ \mathbf{p}_{2,1}\mathbf{x} - u_2\mathbf{p}_{2,3}\mathbf{x} \\ \mathbf{p}_{2,2}\mathbf{x} - v_2\mathbf{p}_{2,3}\mathbf{x} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (8.14)$$

Vztah (8.14) je lineární nehomogenní (uvažme, že $\mathbf{x} = (x, y, z, 1)$) soustavou rovnic pro neznámé hodnoty souřadnic x, y, z . Protože pro tři neznámé jsou k dispozici čtyři rovnice, jde o soustavu předeterminovanou. Předeterminovanosti lze využít ke zmenšení vlivu chyb, které mohou být do výpočtu vneseny nepřesným zadáním či měřením. Hodnoty u_1, v_1, u_2, v_2 souřadnic měřených v obrazech bývají totiž zatíženy šumem, který vzniká nepřesným odečítáním souřadnic v obraze. Do této kategorie nepřesností patří i případ, kdy jsou v digitalizovaném obraze souřadnice měřeny pouze celočíselnými (nikoli reálnými) hodnotami (někdy udáváme souřadnice pixelu, do kterého obraz rekonstruovaného bodu padne, jako celočíselné). Řešení lineárních předeterminovaných systémů lze provést metodou nejmenších čtverců.

Z předchozího výkladu vyplývá, že jsou-li k dispozici projekční matice \mathbf{P}_i , pak lze provést rekonstrukci souřadnic bodů. Postup, v němž jsou projekční matice stanoveny, se obvykle nazývá kalibrací kamer. Ukážeme metodu, jak lze kalibraci provést. Dosaďme opět (8.13) do prvních dvou řádků vztahu (8.12) a provedme transpozici obou takto získaných rovnic. Obdržíme

$$\begin{aligned} \mathbf{x}^\top \mathbf{p}_{i,1}^\top - u_i \mathbf{x}^\top \mathbf{p}_{i,3}^\top &= 0, \\ \mathbf{x}^\top \mathbf{p}_{i,2}^\top - v_i \mathbf{x}^\top \mathbf{p}_{i,3}^\top &= 0. \end{aligned} \quad (8.15)$$

Uspořádejme zatím neznámé prvky projekční matice \mathbf{P}_i do sloupcového vektoru $\mathbf{p}_i = (\mathbf{p}_{i,1}^\top, \mathbf{p}_{i,2}^\top, \mathbf{p}_{i,3}^\top)^\top$. Vztahy (8.15) lze pak přepsat maticově takto

$$\begin{bmatrix} \mathbf{x}^\top & \mathbf{0}^\top & -u_i \mathbf{x}^\top \\ \mathbf{0}^\top & \mathbf{x}^\top & -v_i \mathbf{x}^\top \end{bmatrix} \begin{bmatrix} \mathbf{p}_{i,1}^\top \\ \mathbf{p}_{i,2}^\top \\ \mathbf{p}_{i,3}^\top \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (8.16)$$

Rozepsáno tedy

$$\begin{bmatrix} x & y & z & 1 & 0 & 0 & 0 & 0 & -u_i x & -u_i y & -u_i z & -u_i \\ 0 & 0 & 0 & 0 & x & y & z & 1 & -v_i x & -v_i y & -v_i z & -v_i \end{bmatrix} \mathbf{p}_i = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (8.17)$$

Vidíme, že každý kalibrační bod přispívá k nalezení prvků projekční matice dvěma rovnicemi. Stanovení dvanácti neznámých prvků projekční matice provedeme na základě rovnic poskytnutých dostatečným počtem kalibračních bodů. Ukážeme ale, že projekční matici bude možné pomocí kalibračních bodů stanovit až na reálného násobitele. Uvedenou nejednoznačnost lze snadno ilustrovat pomocí rovnice (8.7). Násobme levou i pravou stranu rovnice libovolným nenulovým reálným násobitelem

μ . Protože v homogenních souřadnicích vektor $\mu \mathbf{u}_i$ reprezentuje tentýž bod obrazu jako vektor \mathbf{u}_i , musí být také matice $\mu \mathbf{P}_i$ nutně správnou projekční maticí. Nejednoznačnost vyřešíme zavedením doplňující podmínky. Snadno lze např. realizovat postup, kdy některý z prvků projekční matice položíme roven zvolené hodnotě (např. roven jedné). Jistou potíž však v tomto případě působí rozhodnutí, který ze členů projekční matice bychom měli takto nastavit (zejména se nesmíme pokoušet takto nastavit člen, který by měl vyjít nulový). Čistší možností je požadovat, aby vhodná norma projekční matice byla rovna zvolené hodnotě. Můžeme např. požadovat, aby součet čtverců všech prvků projekční matice byl roven jedné. V tomto případě však bohužel obdržíme nelineární úlohu, což je podstatná komplikace.

Podkapitolu můžeme uzavřít konstatováním, že máme-li k dispozici pro každou kameru alespoň 6 kalibračních bodů a zvolíme-li doplňující podmínku, můžeme pro řešení prvků projekční matice sestavit systém rovnic a prvky jednoznačně vyřešit. I při použití minimálního počtu šesti bodů je ovšem tento systém předeterminovaný. Pro eliminaci šumu (šumem opět mohou být zatíženy zejména hodnoty souřadnic u_i, v_i změřené v obrazech) bývá však v praxi počet kalibračních bodů mnohem vyšší, než je uvedené teoreticky nutné minimum.

8.4 Relativní kalibrace a rekonstrukce

Často přichází v úvahu úloha, kdy můžeme měřit pouze souřadnice bodů v obrazech získaných kamerami. Trojrozměrné souřadnice žádného bodu ve scéně při tom nejsou k dispozici. Taková situace může nastat např. v případě vizuálního systému robota, který pracuje v předem neznámé scéně. Ukážeme, že i za této situace lze scénu rekonstruovat. Uvažujme opět situaci, kdy máme k dispozici dvě kamery, případně dva obrazy sejmuté ze dvou různých míst jedinou kamerou. Zavedeme matici rotace \mathbf{R} a vektor posunutí $\mathbf{b} = O_2 - O_1$, které provádějí transformaci ze souřadného systému druhé kamery do souřadného systému první kamery podle vztahu

$$\mathbf{x}_1 = \mathbf{R}\mathbf{x}_2 + \mathbf{b}. \quad (8.18)$$

Připomeňme, že \mathbf{R} je ortonormální matice a že $\mathbf{R} = \mathbf{R}_z\mathbf{R}_y\mathbf{R}_x$, kde $\mathbf{R}_x, \mathbf{R}_y, \mathbf{R}_z$ jsou matice odpovídající postupně prováděným rotacím kolem os x_2, y_2 a z_2 . Zavedeme vektor $\varphi = (\varphi_x, \varphi_y, \varphi_z)^\top$ obsahující odpovídající úhly rotace. Matice $\mathbf{R}_x, \mathbf{R}_y, \mathbf{R}_z$

pak mají tvar

$$\mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi_x & \sin \varphi_x \\ 0 & -\sin \varphi_x & \cos \varphi_x \end{bmatrix}, \quad (8.19)$$

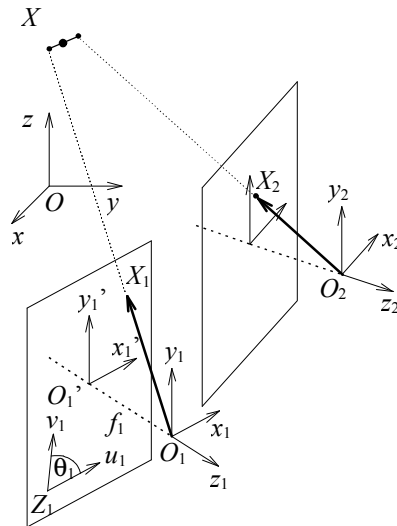
$$\mathbf{R}_y = \begin{bmatrix} \cos \varphi_y & 0 & -\sin \varphi_y \\ 0 & 1 & 0 \\ \sin \varphi_y & 0 & \cos \varphi_y \end{bmatrix}, \quad (8.20)$$

$$\mathbf{R}_z = \begin{bmatrix} \cos \varphi_z & \sin \varphi_z & 0 \\ -\sin \varphi_z & \cos \varphi_z & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (8.21)$$

Jak ukážeme později, lze v procesu relativní kalibrace určit právě matici \mathbf{R} a vektor \mathbf{b} . Známe-li matici \mathbf{R}_1 a vektor \mathbf{o}_1 určující polohu první kamery v globálním souřadném systému scény, pak můžeme hodnoty \mathbf{R}_2 , \mathbf{o}_2 ze vztahu (8.1) určující polohu druhé kamery vyjádřit vztahy

$$\mathbf{o}_2 = \mathbf{R}_1 \mathbf{b} + \mathbf{o}_1, \quad \mathbf{R}_2 = \mathbf{R}_1 \mathbf{R}. \quad (8.22)$$

Hodnoty \mathbf{R}_1 , \mathbf{o}_1 popisují vztah mezi souřadným systémem první kamery a globálním souřadným systémem scény. Tento vztah nelze v procesu relativní kalibrace určit. Někdy mohou být hodnoty \mathbf{R}_1 , \mathbf{o}_1 známy předem. Jestliže tomu tak není, pak pro mnoho úloh postačí položit $\mathbf{R}_1 = \mathbf{I}$ a $\mathbf{o}_1 = \mathbf{0}$, což znamená, že globální souřadný systém scény ztotožníme se souřadným systémem první kamery a souřadnice bodů ve scéně tak měříme v souřadném systému první kamery.



Obr. 8.4: Rekonstrukce polohy bodu

Předpokládejme nyní nejprve, že kalibrace již byla provedena a že matice \mathbf{R}_i a vektor \mathbf{o}_i ($i = 1, 2$) z rovnice (8.1) jsou již pro obě kamery známy. Nejprve ukážeme, jak lze provést rekonstrukci trojrozměrných souřadnic bodu, známe-li souřadnice jeho průmětů v obrazech získaných oběma kamerami. Nechť X označuje rekonstruovaný bod. V souřadném systému (O, x, y, z) nechť je bod X reprezentován vektorem $\mathbf{x} = (x, y, z)$. Nechť X_i je obraz bodu X získaný i -tou kamerou. V souřadném systému (Z_i, u_i, v_i) nechť je bod X_i reprezentován vektorem $\mathbf{u}_i = (u_i, v_i, 1)^\top$. V souřadném systému (O, x, y, z) je přímka $\langle O_i X_i \rangle$ popsána rovnicí $\mathbf{o}_i + \lambda_i \mathbf{R}_i \mathbf{F}_i \mathbf{Q}_i \mathbf{u}_i$, kde λ_i je reálný parametr. Obě přímky $\langle O_i X_i \rangle$ ($i = 1, 2$) by se teoreticky měly protínat v bodě X . Opět je ovšem oprávněné předpokládat, že hodnoty souřadnic v obrazech nelze měřit zcela přesně. Proto nechť v tomto odstavci pojednávajícím o rekonstrukci souřadnic bodu značí \mathbf{u}_i prakticky zjištěné vektory souřadnic, jejichž složky jsou zatíženy šumem měření. Podobně bude nyní X_i značit prakticky zjištěnou polohu průmětu bodu X v obraze získaném i -tou kamerou. Je-li měření zatíženo chybami, pak již bod X obecně nebude ležet současně na obou přímkách $\langle O_i X_i \rangle$ (obr. 8.4). Polohu bodu X lze stanovit na základě podmínky, aby součet čtverců vzdáleností bodu X od obou přímek byl minimální. Hledáme tedy minimum

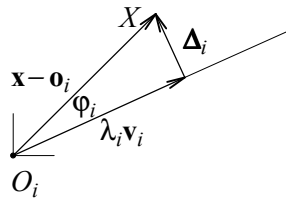
$$\min_{\mathbf{x}} \sum_{i=1}^2 \text{dist}^2(\mathbf{x}, \mathbf{o}_i + \lambda_i \mathbf{R}_i \mathbf{F}_i \mathbf{Q}_i \mathbf{u}_i). \quad (8.23)$$

Pro stručnost matematických výrazů, které budou následovat, zavedeme vektor

$$\mathbf{v}_i = \frac{\mathbf{R}_i \mathbf{F}_i \mathbf{Q}_i \mathbf{u}_i}{|\mathbf{R}_i \mathbf{F}_i \mathbf{Q}_i \mathbf{u}_i|}. \quad (8.24)$$

Na základě vlastností skalárního součinu pro hodnotu parametru λ_i snadno nalezneme vztah

$$\lambda_i = |\mathbf{x} - \mathbf{o}_i| \cos \varphi_i = (\mathbf{x} - \mathbf{o}_i)^\top \mathbf{v}_i. \quad (8.25)$$



Obr. 8.5: Stanovení vzdálenosti bodu X od promítacího paprsku

Nechť Δ_i označuje vektor vzdálenosti bodu X od přímky $\langle O_i X_i \rangle$ (obr. 8.5). Je zřejmé, že $\Delta_i = (\mathbf{x} - \mathbf{o}_i) - \lambda_i \mathbf{v}_i$. Součet čtverců vzdáleností můžeme nyní vyjádřit

pomocí vztahů

$$\begin{aligned}\varepsilon &= \sum_{i=1}^2 \Delta_i^\top \Delta_i = \sum_{i=1}^2 [(\mathbf{x} - \mathbf{o}_i) - \lambda_i \mathbf{v}_i]^\top [(\mathbf{x} - \mathbf{o}_i) - \lambda_i \mathbf{v}_i] = \\ &= \sum_{i=1}^2 [\mathbf{x}^\top \mathbf{x} - 2\mathbf{o}_i^\top \mathbf{x} - 2\lambda_i \mathbf{v}_i^\top \mathbf{x} + \mathbf{o}_i^\top \mathbf{o}_i + 2\lambda_i \mathbf{o}_i^\top \mathbf{v}_i + \lambda_i^2 \mathbf{v}_i^\top \mathbf{v}_i].\end{aligned}\quad (8.26)$$

Derivováním získáme

$$\frac{\partial \varepsilon}{\partial \mathbf{x}} = 2 \sum_{i=1}^2 \left[\mathbf{x} - \mathbf{o}_i - \frac{\partial \lambda_i}{\partial \mathbf{x}} \mathbf{v}_i^\top \mathbf{x} - \lambda_i \mathbf{v}_i + \frac{\partial \lambda_i}{\partial \mathbf{x}} \mathbf{o}_i^\top \mathbf{v}_i + \lambda_i \frac{\partial \lambda_i}{\partial \mathbf{x}} \mathbf{v}_i^\top \mathbf{v}_i \right].\quad (8.27)$$

Z rovnice (8.25) vyplývá, že $\partial \lambda_i / \partial \mathbf{x} = \mathbf{v}_i$. Položíme-li $\partial \varepsilon / \partial \mathbf{x} = \mathbf{0}$, pak po úpravách obdržíme vztah

$$\sum_{i=1}^2 [\mathbf{x} - (\mathbf{v}_i^\top \mathbf{x}) \mathbf{v}_i] = \sum_{i=1}^2 [\mathbf{o}_i - (\mathbf{o}_i^\top \mathbf{v}_i) \mathbf{v}_i].\quad (8.28)$$

Řešením rovnice (8.28) získáme pro hledaný vektor \mathbf{x} předpis

$$\mathbf{x} = \left[\sum_{i=1}^2 (\mathbf{I} - \mathbf{v}_i \mathbf{v}_i^\top) \right]^{-1} \sum_{i=1}^2 [\mathbf{o}_i - (\mathbf{o}_i^\top \mathbf{v}_i) \mathbf{v}_i].\quad (8.29)$$

Nyní se vrátíme k problému kalibrace a ukážeme, jak lze zjistit matici \mathbf{R} a vektor \mathbf{b} , kterých lze podle rovnice (8.22) použít ke zjištění matice \mathbf{R}_2 a vektoru \mathbf{o}_2 (jak jsme již dříve uvedli, předpokládáme, že matici \mathbf{R}_1 a vektor \mathbf{o}_1 známe). Uvažujme bod X scény. V obrazech získaných první a druhou kamerou se tento bod zobrazuje na body X_1, X_2 . V souřadném systému (O_1, x_1, y_1, z_1) jsou směry přímk $\langle O_1 X_1 \rangle, \langle O_2 X_2 \rangle$ reprezentovány vektory $\mathbf{x}_1, \mathbf{R}\mathbf{x}_2$. Jestliže jsou polohy bodů X_1, X_2 v obrazech určeny přesně, pak leží přímky $\langle O_1 X_1 \rangle, \langle O_2 X_2 \rangle, \langle O_1 O_2 \rangle$ v jedné rovině. Podmínku koplanarity lze matematicky zapsat vztahem $\mathbf{x}_1 \cdot (\mathbf{b} \times (\mathbf{R}\mathbf{x}_2)) = 0$. Zápisem $[\mathbf{b}]_\times$ označíme antisymetrickou matici realizující vektorový součin, tj. matici, která pro každý vektor \mathbf{a} vyhovuje podmínce $[\mathbf{b}]_\times \mathbf{a} = \mathbf{b} \times \mathbf{a}$. Lze snadno ověřit, že platí

$$[\mathbf{b}]_\times = \begin{bmatrix} 0 & -b_z & b_y \\ b_z & 0 & -b_x \\ -b_y & b_x & 0 \end{bmatrix}.\quad (8.30)$$

Přepíšeme-li nyní podmínku koplanarity do maticového tvaru, obdržíme rovnici $\mathbf{x}_1^\top [\mathbf{b}]_\times \mathbf{R}\mathbf{x}_2 = 0$. Dosadíme-li do podmínky rovnici (8.4), dostaneme

$$\mathbf{u}_1^\top \mathbf{Q}_1^\top \mathbf{F}_1^\top [\mathbf{b}]_\times \mathbf{R} \mathbf{F}_2 \mathbf{Q}_2 \mathbf{u}_2 = 0,\quad (8.31)$$

případně

$$\mathbf{u}_1^\top \mathbf{Q}_1^\top \mathbf{F}_1^\top [\mathbf{b}]_\times \mathbf{R}_z \mathbf{R}_y \mathbf{R}_x \mathbf{F}_2 \mathbf{Q}_2 \mathbf{u}_2 = 0\quad (8.32)$$

nebo také

$$u_1^\top \mathbf{C} \mathbf{u}_2 = 0, \quad \text{kde } \mathbf{C} = \mathbf{Q}_1^\top \mathbf{F}_1^\top [\mathbf{b}]_\times \mathbf{R} \mathbf{F}_2 \mathbf{Q}_2 = \mathbf{Q}_1^\top \mathbf{F}_1^\top [\mathbf{b}]_\times \mathbf{R}_z \mathbf{R}_y \mathbf{R}_x \mathbf{F}_2 \mathbf{Q}_2. \quad (8.33)$$

Rovnice (8.31), (8.32), (8.33) bývají nazývány rovnicemi koplanarity. Podobně bývá matice \mathbf{C} z rovnice (8.33) nazývána maticí koplanarity. Každá dvojice obrazů jednoho kalibračního bodu popsána souřadnicemi \mathbf{u}_1 v prvním obraze a souřadnicemi \mathbf{u}_2 v druhém obraze dává vzniknout jedné rovnici koplanarity. Je-li k dispozici q kalibračních bodů, pak lze sestavit soustavu q rovnic koplanarity. Této soustavy lze pak použít k řešení neznámých vnějších a vnitřních parametrů kamery, které jsou obsaženy v maticích \mathbf{Q}_1 , \mathbf{Q}_2 , \mathbf{F}_1 , \mathbf{F}_2 , $[\mathbf{b}]_\times$, \mathbf{R} .

Z rovnice (8.31) vyplývá, že řešení bude možné provést až na měřítko vektoru \mathbf{b} . Jestliže je totiž \mathbf{b} vektor, který vyhovuje rovnici (8.31), pak také vektor $\mu \mathbf{b}$, kde μ je reálný násobitel, rovnici vyhovuje. Abychom úlohu specifikovali jednoznačně, zvolíme doplňující podmínku. Můžeme např. požadovat $|\mathbf{b}| = d$, kde d je zvolená hodnota (např. $d = 1$). V případě euklidovské metriky lze pak podmínku rozepsat do tvaru

$$b_x^2 + b_y^2 + b_z^2 = d^2. \quad (8.34)$$

Ze skutečnosti, že délku vektoru \mathbf{b} je zapotřebí zvolit, vyplývá, že rekonstrukci scény bude možné provést až na měřítko. To znamená, že pro různé hodnoty d budeme dostávat podobné, avšak různě veliké scény. Problém lze jednoduše vyřešit tak, že kalibraci provedeme při $|\mathbf{b}| = 1$. Známe-li pak skutečnou vzdálenost mezi některými dvěma body ve scéně, můžeme stanovit měřítko, kterým je třeba dodatečně násobit všechny souřadnice ve scéně (rozměry scény) tak, aby požadované vzdálenosti mezi body bylo skutečně dosaženo.

Z rovnice (8.31) vidíme, že při relativní kalibraci lze stanovit vzájemnou polohu kamer a vnitřní parametry kamery. Vzájemná poloha kamer je popsána maticí \mathbf{R} a vektorem \mathbf{b} . Protože však může existovat nejvýše 7 nezávislých rovnic koplanarity (podrobný důkaz tohoto tvrzení neuvádíme, poněvadž by přesáhl rámec tohoto textu), lze pomocí rovnic koplanarity a s využitím doplňující podmínky (8.34) zjistit nejvýše osm parametrů. To však nemusí být na závadu. Často lze totiž předpokládat, že se některé parametry (zejména vnitřní parametry kamery obsažené v maticích \mathbf{Q}_1 , \mathbf{Q}_2) v čase nemění, a lze je proto předem stanovit předchozím měřením (kalibrací) v laboratorních podmínkách.

Řešení soustavy kalibračních rovnic je obtížné, protože se jedná o nelineární problém. V posledním desetiletí bylo prezentováno několik přístupů k jeho řešení. Zde nastíníme jednu z možných variant. Vyjdeme z rovnic ve tvaru (8.32). Předpokládáme, že máme k dispozici q kalibračních bodů. Souřadnice průmětů i -tého kalibračního bodu v obraze první a druhé kamery označíme postupně $\mathbf{u}_1^{(i)}$, $\mathbf{u}_2^{(i)}$. Každá dvojice $\mathbf{u}_1^{(i)}$, $\mathbf{u}_2^{(i)}$ dává vzniknout jedné rovnici koplanarity. Dostaneme tak celkem q rovnic koplanarity, kterých využijeme spolu s podmínkou (8.34) ke stanovení hodnoty parametrů f_1 , f_2 , φ_x , φ_y , φ_z , b_x , b_y , b_z . Je zřejmé, že nejmenší hodnota

q , při níž lze úlohu řešit, je 7. Prakticky však bývá q vyšší. Předeterminovanosti soustavy lze opět využít k redukci vlivu šumu, kterým může být zatíženo měření souřadnic vektorů $\mathbf{u}_1^{(i)}$, $\mathbf{u}_2^{(i)}$. Hledané hodnoty parametrů uspořádáme do vektoru $\mathbf{p} = (f_1, f_2, \varphi_x, \varphi_y, \varphi_z, b_x, b_y, b_z)^\top$. Pro dvojici obrazů i -tého kalibračního bodu můžeme pak rovnici (8.32) zapsat přehledně ve tvaru

$$F(\mathbf{u}_1^{(i)}, \mathbf{u}_2^{(i)}, \mathbf{p}) = 0. \quad (8.35)$$

Podmínka (8.34) spolu s q ($q > 7$) rovnicemi (8.35) tvoří předeterminovaný nekonzistentní systém (nekonzistence vyplývá z přítomnosti šumu ve vstupních hodnotách). Nekonzistentní systém nemá přísně vzato žádné řešení. Lze pouze najít řešení, které systému v jistém smyslu vyhovuje nejlépe. Můžeme například najít vektor parametrů \mathbf{p} tak, aby součet čtverců reziduí rovnic (8.35) byl minimální. Hledáme tedy takovou hodnotu vektoru parametrů, která minimalizuje funkci

$$\Phi = \sum_{i=1}^q \left[F(\mathbf{u}_1^{(i)}, \mathbf{u}_2^{(i)}, \mathbf{p}) \right]^2, \quad (8.36)$$

a to za současného splnění podmínky (8.34). Hledání minima funkce za doplňujících podmínek lze provést metodou Lagrangeových multiplikátorů (jedná se o standardní matematickou metodu). Metoda převádí problém na problém nalezení prostého minima (bez doplňující podmínky) nové funkce, která však má větší počet argumentů. V našem případě je jistou komplikací, že je úloha nelineární. Nelineární úlohu minimalizace lze řešit např. gradientní metodou.

8.5 Epipolára

Uvažujme bod X scény. V obrazech získaných první a druhou kamerou jsou souřadnice průmětů tohoto bodu popsány vektory $\mathbf{u}_1 = (u_1, v_1, 1)$, $\mathbf{u}_2 = (u_2, v_2, 1)$. Dosadíme-li do rovnice (8.18) rovnici (8.4) zapsanou pro první a druhou kameru, snadno získáme vztah

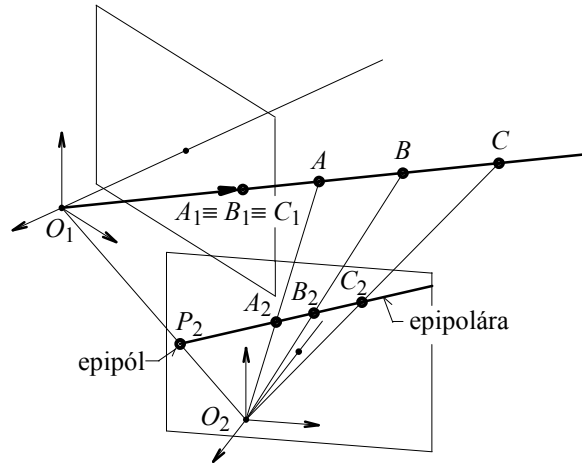
$$\lambda_1 \mathbf{F}_1 \mathbf{Q}_1 \mathbf{u}_1 = \lambda_2 \mathbf{R} \mathbf{F}_2 \mathbf{Q}_2 \mathbf{u}_2 + \mathbf{b}. \quad (8.37)$$

Odtud máme

$$\lambda_2 \mathbf{u}_2 = \lambda_1 \mathbf{Q}_2^{-1} \mathbf{F}_2^{-1} \mathbf{R}^\top \mathbf{F}_1 \mathbf{Q}_1 \mathbf{u}_1 - \mathbf{Q}_2^{-1} \mathbf{F}_2^{-1} \mathbf{R}^\top \mathbf{b}. \quad (8.38)$$

Součin $\lambda_2 \mathbf{u}_2$ ve vztahu (8.38) reprezentuje v homogenních souřadnicích průmět bodu X v obraze druhé kamery. Rovnice (8.38) je homogenní parametrickou rovnicí přímky, v níž je parametrem hodnota λ_1 . Můžeme tedy učinit tento závěr: Jestliže máme v obraze první kamery bod, jehož souřadnice jsou $\mathbf{u}_1 = (u_1, v_1, 1)$, pak poloha jemu odpovídajícího bodu v obraze druhé kamery závisí na hodnotě λ_1 (použitím vztahu (8.4) lze snadno ověřit, že platí $\lambda_1 = -z_1/f_1$). Všechny možné polohy vytvoří v obraze získaném druhou kamerou přímku. Tato přímka se nazývá epipolárou. Pro

$\lambda_1 = 0$ rovnice (8.38) určuje v obraze získaném druhou kamerou polohu průmětu středu projekce první kamery. Tento bod se nazývá epipól. Epipolára i epipól jsou vyobrazeny na obr. 8.6. Poznamenejme, že odvozené vztahy je samozřejmě možné také obrátit, což znamená, že k bodu v obraze získaném druhou kamerou je možné nalézt epipoláru v obraze získaném první kamerou. Totéž platí i pro epipól.



Obr. 8.6: Epipolára a epipól

V případě kamer s rovnoběžnými osami (podkapitola 8.2) se rovnice (8.38) podstatně zjednoduší. V tomto případě totiž máme $\mathbf{F}_1 = \mathbf{F}_2$ (proto dále píšeme jen \mathbf{F}), $\mathbf{Q}_1 = \mathbf{Q}_2 = \mathbf{I}$ (jednotková matice) a $\mathbf{R} = \mathbf{I}$. Z rovnice (8.38) obdržíme vztah $\lambda_2 \mathbf{u}_2 = \lambda_1 \mathbf{u}_1 - \mathbf{F}^{-1} \mathbf{b}$. Uvážíme-li, že $\mathbf{b} = (b_x, 0, 0)$, pak z rovnice pro souřadnici z zjišťujeme, že $\lambda_2 = \lambda_1$ (dále proto píšeme jen λ). Parametrické rovnice epipoláry pak jsou $u_2 = u_1 - b_x/\lambda$, $v_2 = v_1$.

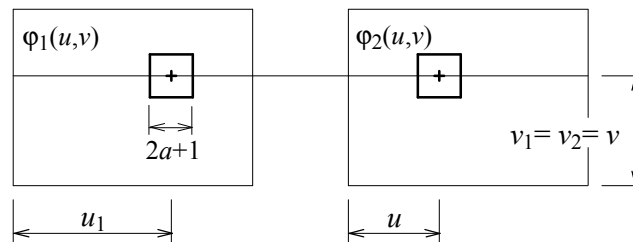
8.6 Automatizované hledání korespondence

Pro rekonstrukci polohy bodu je zapotřebí identifikovat jeho průměty v obraze získaném první a druhou kamerou. Zdá se, že je možné uvažovat i o automatizaci této úlohy. Lze postupovat např. v následujících krocích: 1) V obraze získaném první kamerou identifikujeme body zájmu např. některým z detektorů rohů. 2) Pro každý bod zájmu nalezený v prvním obraze určíme korespondující bod v druhém obraze. Druhý krok popíšeme podrobněji. Necht X_1 je bod zájmu nalezený v prvním obraze a jeho souřadnice necht jsou u_1, v_1 . Z předchozí podkapitoly víme, že korespondující bod X_2 bude v druhém obraze ležet na epipoláře. Předpokládáme, že kromě hodnoty λ_1 , která je parametrem, jsou všechny zbývající hodnoty na pravé straně rovnice (8.38) k dispozici (řekněme, že již byla provedena kalibrace). Můžeme proto stanovit rovnici epipoláry, na níž bod X_2 leží. Přesnou polohu bodu X_2 nalezneme tak, že

systematicky prověřujeme všechny body na epipoláře (při praktické implementaci můžeme postupovat např. po celých pixelech). Označme $\varphi_1(u, v)$, $\varphi_2(u, v)$ obrazové funkce prvního a druhého obrazu. Necht u, v jsou souřadnice prověřovaného bodu ve druhém obraze (u_1, v_1 jsou souřadnice v obraze prvním, tedy souřadnice bodu X_1). Otázku, zda prověřovaný bod koresponduje s X_1 , můžeme rozhodnout na základě porovnání okolí bodu X_1 s okolím bodu prověřovaného. K porovnání lze použít např. vztahu (obr. 8.7)

$$\text{Diff}(u, v) = \sum_{\Delta u=-a}^a \sum_{\Delta v=-a}^a [\varphi_1(u_1 + \Delta u, v_1 + \Delta v) - \varphi_2(u + \Delta u, v + \Delta v)]^2. \quad (8.39)$$

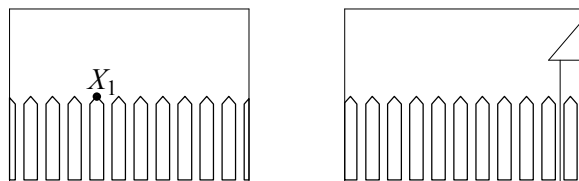
Je zřejmé, že vztah (8.39) porovnává okolí čtvercového tvaru o straně $2a + 1$. Čím menší hodnota vyjde, tím jsou si okolí podobnější. Za bod korespondující s X_1 lze považovat takový bod X_2 , který splňuje následující kriteria: 1) hodnota vypočítaná podle vztahu (8.39) je pro X_2 minimální a menší než nějaký zvolený práh, 2) pro všechny ostatní body na epipoláře je hodnota (8.39) významně větší než pro bod X_2 . Druhé kritérium zohledňuje požadavek jednoznačnosti detekce.



Obr. 8.7: Hledání korespondence porovnáním okolí bodu (kamery s rovnoběžnými optickými osami)

Ačkoli právě popsaný postup vypadá dosti přímočaře, skrývá mnohá úskalí. Prvním problémem je volba vhodné velikosti okénka. Je-li okénko malé, hrozí, že detekce nebude jednoznačná. Při velkém okénku se naopak významněji uplatní fakt, že obrazy sejmuté z různých míst jsou přesně vzato různé a nelze obecně očekávat shodu příliš velkých oblastí ve smyslu vztahu (8.39). Vztah (8.39) dále předpokládá, že okénku velikosti $2a + 1$ v prvním obraze odpovídá ve druhém obraze okénko téže velikosti a že hrany okének jsou v obou obrazech rovnoběžné se souřadnými osami. To je splněno pouze ve speciálních případech (např. v případě kamer s rovnoběžnými optickými osami z podkapitoly 8.2).

Otázka jednoznačnosti detekce korespondence je klíčová. Přísně vzato je zapotřebí se smířit s faktem, že mohou existovat případy, kdy jednoznačnou detekci zajistit nelze. Uvažme např. příklad dle obr. 8.8. Ani člověku se zde bez doplňujících informací nepodaří stanovit, který bod v druhém obraze koresponduje s bodem X_1 . Tím méně lze u podobných úloh očekávat řešení automatizované. Automatizované



Obr. 8.8: K nejednoznačnosti řešení problému korespondence

hledání korespondence je z výše uvedených důvodů úlohou, jejíž obecné řešení je obtížné. I když bylo pro automatizované hledání korespondence publikováno více postupů, nelze zatím žádný z nich považovat za zcela uspokojivý. Ve zbytku této podkapitoly podáme přehled některých základních přístupů, které s v průběhu řešení problému hledání korespondence objevily.

Na stereokorespondenci můžeme pohlížet jako na proces hledání podobnosti mezi dvěma skupinami dat v obrazech. Nalezení správné stereokorespondence je klíčovým článkem pro následnou 3D rekonstrukci scény. Zatímco náš mozek je schopen provádět tyto operace zcela automaticky, bez přispění vědomí, pro počítač je tento problém obtížně uchopitelný. Odpovídající algoritmy se vyvíjejí více než 30 let a stále se nedá říci, že by se podařilo najít optimální řešení.

Metody hledání korespondence se dělí podle typu výsledné disparitní mapy na husté (obsahující hodnotu disparity pro téměř všechny body v obraze) a na řídké (disparita je přiřazena jen význačným bodům, např. rohům). Další možností, jak tyto algoritmy rozčlenit, je na základě přístupu k hledání korespondence. Lokální metody hledají korespondenci jen na základě informací získaných z konečného okolí hledaného bodu, nejčastěji se jedná o malá čtvercová okénka. Globální metody přistupují k problému korespondence jako k optimalizační úloze, kdy se určí disparita a zakryty na celé ploše obrazu. V následujících podkapitolách si ukážeme příklady obou přístupů.

8.6.1 Okénkové algoritmy

Okénkové algoritmy vycházejí z předpokladu podobnosti okolí korespondujících bodů. Na prvním obraze vymežíme malou oblast, nejčastěji čtvercového tvaru, a tu postupně porovnáváme s potenciálními pozicemi, kde by se mohl nalézat korespondující bod druhého obrazu. Epipolární geometrie nám zajišťuje výskyt korespondujícího bodu na epipoláře, u rektifikovaných obrazů dokonce na stejné horizontální souřadnici. Pro porovnání oblastí okolí bodu se nejčastěji používají tři druhy metrik: rozdíl hodnot jasu (SAD, SSD), korelace (NCC) a metrika tzv. rank transformace. Máme-li definováno okénko $u \times v$, jsme schopni určit míru podobnosti pro danou hodnotu

disparity d pro jednotlivé metriky pomocí vztahů

$$\text{SAD}(u, v) = \sum_{u, v} | \varphi_1(u, v) - \varphi_2(u + d, v) |, \quad (8.40)$$

$$\text{SSD}(u, v) = \sum_{u, v} (\varphi_1(u, v) - \varphi_2(u + d, v))^2, \quad (8.41)$$

$$\text{NCC}(u, v) = \sum_{u, v} \frac{(\varphi_1(u, v) - \bar{\varphi}_1) \cdot (\varphi_2(u + d, v) - \bar{\varphi}_2)}{\sqrt{(\varphi_1(u, v) - \bar{\varphi}_1)^2 \cdot (\varphi_2(u + d, v) - \bar{\varphi}_2)^2}}. \quad (8.42)$$

Algoritmy využívající předchozí metriky jsou jednoduše implementovatelné a také snadno paralelizovatelné. Hlavní slabina těchto metod ale tkví ve volbě velikosti použitého okna. Malá okénka mají menší diskriminační schopnost a budou vytvářet větší počet chybných korespondencí. Oproti tomu velká okénka mají sice větší odolnost proti šumu a poruchám v obraze, ale současně také rozmazávají disparitní mapu, protože nejsou schopny zachytit dostatečně jemně změnu disparity. Na obrázku 8.9 je výsledná disparitní mapa při použití metriky SAD. Z obrázku je patrné, že lokální metody selhávají na velkých jednolitých plochách. Problém volby velikosti porovnávaného okénka se snaží řešit algoritmy mající adaptivní velikost okna, jehož velikost a tvar se mění podle potřeby, nejčastěji na základě členitosti disparitní mapy a průběhu jasové funkce porovnávaných obrazů.

Kromě klasických metod využívajících přímo jasové hodnoty pixelů existují i alternativní metody využívající neparametrické transformace vstupních obrazů. Příkladem je tzv. rank transformace. Rank transformace udává počet bodů v malém okénku kolem centrálního bodu, jejichž jasová hodnota je menší než jasová hodnota středového bodu. Výsledná hodnota tak závisí pouze na relativním uspořádání jasových hodnot, ne na hodnotách samotných. Poté, co je provedena tato transformace, je možné bloky porovnávat běžnou L_2 normou.

$$\text{Rank}(u, v) = \sum_{u, v} (\varphi'_1(u, v) - \varphi'_2(u + d, v))^2, \quad (8.43)$$

$$\varphi'_k(u, v) = \sum_{m, n} R(m, n, u, v), \quad (8.44)$$

$$R(m, n, u, v) = \begin{cases} 0 & \varphi_k(m, n) \geq \varphi_k(u, v) \\ 1 & \varphi_k(m, n) < \varphi_k(u, v) \end{cases}. \quad (8.45)$$

Výhodou rank transformace je větší odolnost proti šumu. Z důvodu ztráty části informací je však snížena i rozlišovací schopnost porovnávání.

8.6.2 Dynamické programování

Dynamické programování je technika, která se často užívá pro řešení optimalizačních úloh. Tuto techniku lze použít tehdy, lze-li původní úlohu rozložit na menší

podúlohy, z jejichž optimálních řešení se poté složí řešení celé původní úlohy. Základem je Bellmanův princip optimality, který říká, že podstrategie optimální strategie je opět optimální. Využití dynamického programování vede k časově efektivnějším algoritmům, zpravidla však za cenu mírného zvýšení paměťových nároků.

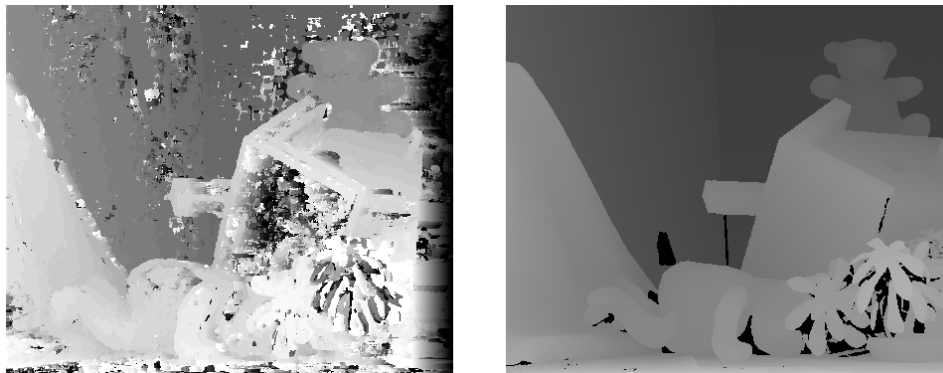
Předtím, než se podíváme na princip této metody, je třeba zavést některá omezení, bez kterých by tato metoda nebyla použitelná. Při hledání korespondence budeme předpokládat, že bod X_1 může korespondovat pouze s jedním bodem X_2 druhého obrazu. Zanedbáváme tím situace, ke kterým může docházet v případech, kdy se objekt snímáný jednou kamerou promítne do podoby jednoho pixelu, zatímco druhá kamera jej vlivem natočení vidí jako pixely dva. Dále musíme předpokládat, že změny v obrazech způsobené rozdílným umístěním kamer jsou natolik malé, že nedochází k záměně pořadí bodů ležících na epipoláře. Nachází-li se dva body A a B (oba leží na epipoláře) v těsné blízkosti, a bod A je situován vlevo od bodu B , předpokládáme, že na druhém snímku bude bod A ležet opět vlevo od bodu B . Jako poslední omezení se často klade požadavek, aby výsledná disparitní mapa byla hladká, tj. disparita sousedních bodů se příliš neodlišovala.

Za předpokladu splnění předchozích podmínek jsme nyní schopni problém korespondence převést na problém nalezení optimální cesty v síti. Hledání korespondence probíhá po jednotlivých řádcích. Podívejme se na příklad na obr. 8.10. Vidíme zde řádky levého a pravého obrazu a vyznačené korespondence mezi jednotlivými body. Barevně odlišeny jsou ty body, ke kterým nelze najít korespondující bod z důvodu zákrytu. Stejnou situaci ilustruje i síť zobrazená v pravé části obrázku.

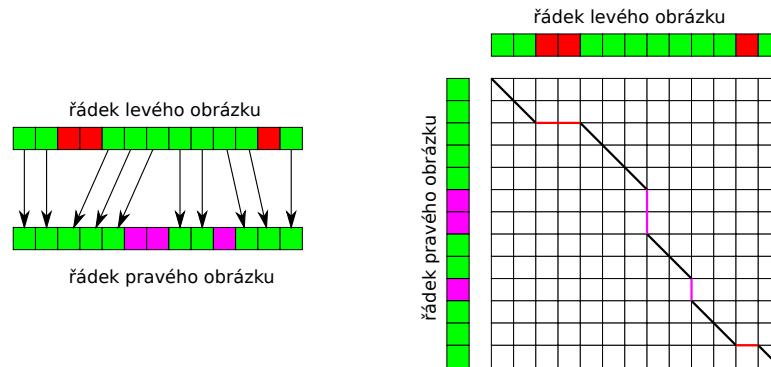
Algoritmus využívající dynamického programování pak funguje následovně: Pro každý řádek levého (I_l) a pravého (I_r) obrazu vypočítáme matici M_h velikosti $W \times W$, kde hodnota W odpovídá šířce vstupních obrazů. Matice M_h je maticí cen korespondence mezi pixely $p(x, y)$ a $q(x + d, y)$ a hodnotou disparity d , přičemž $x + d < W$. Cena M_h je definována jako

$$M_h(i, j) = \min(\text{Diff}(i, j) + M_h(i - 1, j - 1), \lambda + M_h(i - 1, j), \lambda + M_h(i, j - 1)), \quad (8.46)$$

kde λ je hodnota penalizace, která se započítává, pokud není nalezena korespon-



Obr. 8.9: Výsledná disparitní mapa SAD a správné řešení



Obr. 8.10: Optimální cesta v síti

dence, a $\text{Diff}(i, j)$ je funkce porovnávající body, například pomocí podobnosti jejich okolí (SAD, SSD). Implementace výpočtu matice M_h je ilustrována na výpisu 8.6.2.

```

for (int i = 1; i < W; i++)
  for (int j = 1; j < W; j++) {
    min1 = diff(i, j) + M[j-1, i-1 ];
    min2 = lambda + M[j-1, i ];
    min3 = lambda + M[j, i-1 ];
    M[i, j] = min(min1, min2, min3);
    D[i, j] = argmin(min1, min2, min3);
  }

```

Zdrojový kód 1: DP – Výpočet matice M_h

Disparitní mapu pro daný řádek získáme zpětným průchodem matice M_h . Zpětný průchod začíná na pozici $M_h(W, W)$ a pokračuje po směru nejlevnější cesty. Pro implementaci je vhodné si zaznamenávat kromě ceny i směr, kterým jsme se do dané buňky dostali (v kódu matice D). Hodnotu disparity určíme během zpětného průchodu jako rozdíl souřadnic korespondujících bodů $d = i - j$. Výpis 8.6.2 obsahuje možnou implementaci zpětného průchodu matice M_h na základě směru určeného pomocí matice D .

```

i = j = W;
while ( (i != 0) && (j != 0) ) {
  switch (D[i, j]) {
    case 1: i--; j--; break;
    case 2: i--; break;
    case 3: j--; break;
  }
}

```

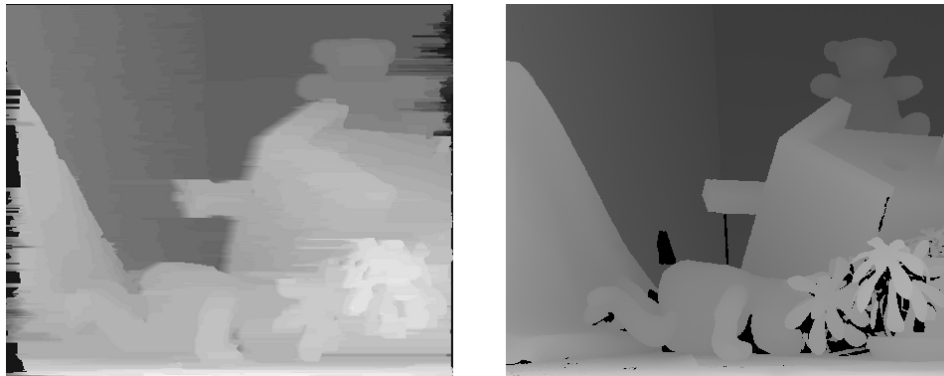
```

    }
    d = i - j;
}

```

Zdrojový kód 2: DP – Zpětný průchod

Dynamické programování poskytuje lepší výsledky než běžné okénkové metody díky minimalizaci ceny korespondence. Zpracování každého řádku obrazu je nezávislé, a proto lze takto snadno algoritmus paralelizovat. Jednou z nevýhod této metody je, že v základním provedení se provádí optimalizace jen v rámci jednoho řádku a nebere se ohled na okolní výsledky. Ukázka výsledné disparitní mapy je na obr. 8.11.



Obr. 8.11: Výsledná disparitní mapa DP a správné řešení

8.6.3 Globální metody

Globální metody hledají minimum funkce ceny ohodnocení disparity celého obrazu. V kontextu globálních metod se berou nejčastěji v úvahu tyto dvě podmínky:

- datová podmínka: Dva body mohou korespondovat jen tehdy, mají-li podobnou intenzitu všech složek.
- požadavek hladkosti: Výsledná disparitní mapa by měla být po částech hladká, tedy dva sousední body musí mít pokud možno co nejvíce podobné hodnoty disparity.

Na základě těchto dvou podmínek je konstruována funkce reprezentující celkovou energii disparit. Algoritmy hledají takovou disparitní funkci d , která minimalizuje funkcional $E(d)$ daný vztahem

$$E(d) = E_d(d) + \lambda E_s(d), \quad (8.47)$$

kde $E_d(d)$ reprezentuje datovou podmínku, $E_s(d)$ předpoklad hladkosti a konstanta λ udává váhu hladkosti. Hodnota $E_d(d)$ udává podobnost dvou pixelů, kterou je možné vypočítat pomocí dříve zmíněných metrik (např. SSD, SAD). Funkce $E_s(d)$ pak může vypadat následovně:

$$E_s(d) = \sum_{u,v} \rho(d(u,v) - d(u+1,v)) + \rho(d(u,v) - d(u,v+1)), \quad (8.48)$$

kde ρ je monotónně rostoucí funkce rozdílu disparit sousedních pixelů. Platí, že čím více se liší hodnota disparit sousedních pixelů, tím větší je i hodnota $E_s(d)$.

Ukázalo se bohužel, že minimalizace funkcionálu $E(d)$ je NP-obtížný problém. Z tohoto důvodu se pro nalezení řešení užívají různé přibližné metody, jakými jsou grafové řezy, propagace zpráv, simulované žíhání a další. Není bohužel v možnostech tohoto textu popsat tyto metody podrobněji.

Literatura

- [1] C.M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006, (ISBN 0-387-31073-8, 978-0-387-31073-2).
- [2] Al Bovik, *Handbook of Image and Video processing 2nd Edition*. Elsevier Academic Press, 2005, (ISBN 0121197921, 978-0121197926).
- [3] B. Budínský, *Analytická a diferenciální geometrie*. SNTL Praha, 1983.
- [4] J.F. Canny, *Finding edges and lines in images*, Technical report AI-TR-720. MIT Artificial Intelligence Laboratory, 1983.
- [5] J.F. Canny. *A computational approach to edge detection*. IEEE PAMI **8**, 1986, 679–698.
- [6] K.R. Castleman. *Digital Image Processing*. Prentice-Hall, 1996, (ISBN 0-13-211467-4).
- [7] R.E. Dougherty. *An Introduction to Morphological Image Processing*. SPIE Press, Bellingham, WA, 1992.
- [8] R.O. Duda, P.E. Hart. *Use of Hough transformation to detect lines and curves in pictures*. Comm. ACM **15**, 1972, 11–15.
- [9] D.E. Dudgeon, R.M. Mersereau. *Multidimensional Digital Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1984.
- [10] R.C. Gonzales and R.E. Woods, *Digital Image Processing, (3rd edition)*. Prentice Hall, 2008 (ISBN 0-13-505267-X).
- [11] A. Hornberg, *Handbook of Machine Vision*. Wiley VCH, 2006, (ISBN 978-3-527-40584-8).
- [12] D. Maar, E. Hildreth. *Theory of edge detection*. Proc. Roy. Soc. London B 207, 1980, 187–217.
- [13] M.A. Penna, R.R. Patterson. *Projective Geometry and its Applications to Computer Graphics*. Prentice-Hall, Englewood Cliffs, NJ, 1986. ISBN 0-13-730649-0.

-
- [14] W.K. Pratt, *Digital image processing: PIKS inside (3rd edition)*. John Wiley and Sons, 2001 (ISBN 0-471-37407-5).
- [15] J. Serra. *Image Analysis and Mathematical Morphology, Vol.1*. Academic Press, New York, 1982.
- [16] J. Serra (ed.). *Image Analysis and Mathematical Morphology, Vol.2*. Academic Press, New York, 1988.
- [17] E. Sojka, M. Němec, T. Fabián. *Matematické základy počítačové grafiky*. VŠB-TU Ostrava, 2011.

Rejstřík

- aliasing, 47
- autokorelace, 34
- autokovariance, 35
- bodový zdroj, 11
- báze, 2
- dilatace, 108
- Diracův impulz, 9
- energetické spektrum signálu, 19
- ergodicita, 36
- eroze, 108
- filtr
 - inverzní, 67
 - rekonstrukční, 44
 - Wienerův, 67
- funkce
 - distribuční, 34
 - Kroneckerova, 4
 - kvantovací, 54
 - vzorkovací, 41
- histogram
 - bimodální, 103
 - jasu, 70
- hrana, 84
- hranice, 84
- hustotu pravděpodobnosti, 34
- kvantování, 54
- křížová korelace, 35
- křížová kovariance, 35
- náhodná proměnná, 33
- náhodné pole, 33
- náhodný proces, 33
- obraz
 - rekonstruovaný, 43
 - vzorkovaný, 43
- operátor
 - Kirschův, 88
 - Prewittové, 87
 - Sobelův, 88
- otevření, 109
- prahování, 103
- princip superpozice, 2
- prostor
 - signálový, 2
 - unitární, 4
- signál, 1
 - diskrétní, 1
 - reprezentace, 2
 - spojitý, 1
- skalární součin, 4
- souřadnice
 - bodů, 1
 - homogenní, 78
- střední hodnota, 34
- teorém
 - Konvoluční, 25
 - Korelační, 26
 - Parsevalův, 26
- transformace
 - afinní, 77
 - Fourierova, 18
 - Houghova, 92
 - separabilní, 18
- uzávěr, 109
- vzorkování, 41